

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2001, 23:48:11 ; Search time 1789.67 seconds
(without alignments)
6231.369 Million cell updates/sec

Title: US-09-735-712-9
676
Sequence: 1 tcatctccttcaattatc.....tgyttaataaaaaaa 676

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_da:*
2: gb_hg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	676	100.0	676	6	AX179789	AX179789 Sequence
2	667.2	98.7	694	9	AF237907	AF237907 Homo sapi
3	666.6	98.6	697	9	AF231127	AF231127 Homo sapi
4	665.6	98.5	691	9	AB013103	AB013103 Homo sapi
5	603	89.2	603	6	AX179781	AX179781 Sequence
6	347.4	51.4	450	6	AX179787	AX179787 Sequence
7	341.8	50.6	417	6	AX179785	AX179785 Sequence
8	246	36.4	246	6	AX179783	AX179783 Sequence
9	185.4	27.4	138097	2	AC027787	AC027787 Homo sapi
10	185.4	27.4	138097	2	AC027787	AC027787 Homo sapi
11	185.4	27.4	147788	2	AC015840	AC015840 Homo sapi
12	185.4	27.4	161039	2	AP003127	AP003127 Homo sapi
13	185.4	27.4	166804	2	AP001034	AP001034 Homo sapi
14	185.4	27.4	166820	2	AC090401	AC090401 Homo sapi
15	157	23.2	161039	2	AP003127	AP003127 Homo sapi
16	148.6	22.0	166804	2	AP001034	AP001034 Homo sapi
17	88	13.0	1076	9	AF068288	AF068288 Homo sapi
18	86.4	12.8	720	9	AB022821	AB022821 Homo sapi
19	86.4	12.8	916	9	AB013102	AB013102 Homo sapi
20	86.4	12.8	1619	9	AF237912	AF237912 Homo sapi
21	86.4	12.8	1669	6	AR035695	AR035695 Sequence
22	86.4	12.8	1669	6	AR083580	AR083580 Sequence
23	76.2	11.3	2805	9	AF367473	AF367473 Homo sapi
24	63.2	9.3	1546	9	HUMIERB	L35848 Homo sapien
25	63.2	9.3	1661	6	A68627	A68627 Sequence 3
26	63.2	9.3	1661	6	AR082014	AR082014 Sequence
27	63.2	9.3	1661	6	AR082014	AR082014 Sequence
28	63.2	9.3	1661	6	AR082014	AR082014 Sequence
29	62.4	9.2	1222	4	SSR236932	AJ236932 Sus scrofa
30	58.6	8.7	708	6	AX101306	AX101306 Sequence
31	58.6	8.7	1010	10	MUSFCERB	J05019 Mouse mast
32	55.6	8.2	899	10	AF237910	AF237910 Mus muscu
33	55.4	8.2	1163	6	AF237915	AF237915 Mus muscu
34	55.4	8.1	1474	6	I07272	I07272 Sequence 19
35	55	8.1	1474	9	HSCD20	X12530 Human mRNA
36	55	8.1	1476	6	I07277	I07277 Sequence 24
37	55	8.1	1482	9	BC002807	BC002807 Homo sapi
38	55	8.1	1597	9	HSCD20A	X07203 Human mRNA
39	54.8	8.1	518	10	AF280401	AF280401 Mus muscu
40	53	7.8	1480	10	AF237914	AF237914 Mus muscu
41	52.6	7.8	862	9	AB013104	AB013104 Homo sapi
42	52.6	7.8	930	9	AF286866	AF286866 Homo sapi
43	52.6	7.8	1131	9	AF212240	AF212240 Homo sapi
44	52.6	7.8	1152	9	AF237908	AF237908 Homo sapi
45	52.6	7.8	1289	9	AF142409	AF142409 Homo sapi

ALIGNMENTS-

RESULT 1
AX179789
LOCUS AX179789 676 bp DNA
DEFINITION Sequence 9 from Patent WO0146417.
ACCESSION AX179789
VERSION AX179789.1 GI:15132148
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 676)
Walker, D.W. and Turner, C.A.
TITLE
Human membrane proteins and polynucleotides encoding the same
having homology to cd20 proteins and ige receptors
JOURNAL
Patent: WO 0146417-A 9 28-JUN-2001;
Genetics Incorporated (US)
FEATURES
Location/Qualifiers
1..676
/organism="Homo sapiens"


```

QY 661 ttaaaataaaaa 672
Db 683 TTTAAATVAAAAA 694

RESULT 3
AF321127 697 bp mRNA PRI 07-FEB-2001
LOCUS Homo sapiens testis-expressed transmembrane-4 protein (TMEM4) mRNA,
DEFINITION complete cds.
ACCESSION AF321127
VERSION AF321127.1 GI:12698681
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS Hulett,M.D., Pagler,E., Hornby,J.R., Hogarth,P.M., Eyre,H.J.,
Baker,E., Crawford,J., Sutherland,G.R., Ohms,S.J. and Parish,C.R.
TITLE Isolation, tissue distribution, and chromosomal localization of a
novel testis-specific human four-transmembrane gene related to CD20
and Fcpsi10nR1-beta
JOURNAL Biochem. Biophys. Res. Commun. 280 (1), 374-379 (2001)
MEDLINE 21092614
REFERENCE 2 (bases 1 to 697)
AUTHORS Hulett,M.D.
TITLE Direct Submission
SUBMITTED (14-NOV-2000) Division of Immunology and Cell Biology,
John Curtin School of Medical Research, Mills Road, Canberra, ACT
JOURNAL 2601, Australia
FEATURES
source 1. 697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12"
1. 697
/gene="TMEM4"
57. 659
/gene="TMEM4"
/codon_start=1
/product="testis-expressed transmembrane-4 protein"
/db_xref="GI:12698682"
/translation="MDSSTASHPVLPVPPETASFEVSTELSTSTFSTQSLQFLA
RKMKLGIQLIFGIMTFSFGVLEFLTKPRPRPFLSGYPWGSVLFNLSGATL
IAVKRTETLILSRIMNPLSALGAIIGIILTPGFLIDONYICGYSHONSQKRAVT
VLLGLITLMTFSIIEFLISLPSILGCHSDCEQCC"
gene
CDS
ORIGIN
BASE COUNT 199 a 142 c 119 g 237 t
Query Match 98.6%; Score 666.6; DB 9; Length 697;
Best Local Similarity 99.4%; Pred. No. 1.5e-145;
Matches 669; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 tatctcccttcaaatatcaccggaacacatcatggtatcgaagcaccgacacagtcagg 60
Db 25 TCATCTCCCTTCAAAATATACACGACCATCATGATGATCAAGCACACGACAGTCGCG 84
QY 61 tgtttcgtattctccccaagaatactgcttcagaatgatgagtcacagaaacttcag 120
Db 85 TGTTCGTGTAATTTCTCCAGAAATACATGCTTCAGAAATATAGTCCACAGAACTTCAG 144
QY 121 ccaagacctttcaactcaaaagcccttcgaaaaattatctgtagaanaatgaaatct 180
Db 145 CCACGACCTTTTCAACTCAAAAGCCCTTGCAAAATATTATTCGTAGAAATGAAATCT 204
QY 181 taggaactatccagatcccggttggaatatgacctttctttggagttatcttctt 240
Db 205 TAGGAGACTATCCAGATCCGTTTGGAATATGACCTTTCTTTGGAGTATCTCCCTT 264

```

```

QY 241 tcaactgttaaaacccatccaaaggttcccttattatcttcttcaagatccattct 300
Db 265 TCACCTTGTAAAACCATATCCAAAGCTTCCCTTATATTTCTTTCAGGATATCCATCT 324
QY 301 ggggctctgtttgttccatcaatctcgaagccttcctaattgagtgagtgaaaaaacc 360
Db 325 GGGGCTCTGTTTGTTCATTAATTCGAGCCCTTCCATTAATGCAAGTAAAGAAAAACA 384
QY 361 cagaactctgataatatgagccgaataatgatacttctttagtgcctcgaagcaatag 420
Db 385 CAGAAACTCTGATTAATATGAGCCGGAATTAATGATTTCTTAATGCTGGAGCAATAG 444
QY 421 ctggaatcaatccctccacattggtttcctcctagatcaaaactacattggtatt 480
Db 445 CTGGAATCATCTTCCTCCACATTGGTTTCATCTAGATCAAAACTACATTGGTTATT 504
QY 481 ctaccaaaatagtcagtgtaagctgttactgctgtctctcttggaatttgatcat 540
Db 505 CTCACCAAAATAGTCAGTGAAGCGTGTACTGCTGCTGTTGGGAATTTGATTACAT 564
QY 541 tgatgacttcagcatatggaattatcatctctctgctctcctaatttgggtgc 600
Db 565 TGATGACTTTCACCATTAATGGAATTAATTCATTTCTGCTGCTTCTCAATTTGGGGTCC 624
QY 601 actcagaagatttgatttggaacaaatggttgactagacactgtagaataagaatgt 660
Db 625 ACTCAGAGGATGTGATGATGGAACAATGTTGTGACTAGCAGTGTGAATTAAGATGTG 684
QY 661 ttaaaataaaaa 673
Db 685 TTTAAATVAAAAA 697

```

```

RESULT 4
AB013103
LOCUS AB013103 691 bp mRNA PRI 20-MAR-2001
DEFINITION Homo sapiens mRNA for MS4A5, complete cds.
ACCESSION AB013103
VERSION AB013103.1 GI:11559213
KEYWORDS MS4A5; CD20-like 2.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Ishibashi,K., Suzuki,M., Sasaki,S. and Imai,M.
TITLE Identification of a new multigene four-transmembrane family (MS4A)
related to CD20, HTm4 and beta subunit of the high-affinity IGE
receptor
JOURNAL Gene 264 (1), 87-93 (2001)
MEDLINE 21142397
REFERENCE 2 (sites)
AUTHORS Ishibashi,K., Sasaki,S. and Marumo,F.
TITLE Cloning of three CD20 homolog from human, putative calcium channels
Unpublished
JOURNAL 3 (bases 1 to 691)
REFERENCE Ishibashi,K.
AUTHORS Direct Submission
TITLE Submitted (20-APR-1998) Kenichi Ishibashi, Tokyo Medical and Dental
University, 2nd Internal Medicine; Yushima 1-5-45, Bunkyo, Tokyo
113-8519, Japan (E-mail:Kishibashi.med2med.tmd.ac.jp,
Tel:81-3-5803-5223, Fax:81-3-5803-0132)
FEATURES
source 1. 691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="testis"
52. 654
/gene="MS4A5"
52. 654
/gene="MS4A5"
52. 654
/gene="MS4A5"
/codon_start=1
/product="MS4A5"

```

/protein_id="BAB18739.1"
/db_xref="GI:1153214"
translation="MDSSTAHPEVLPPEPTATSEVSTELSATSTSPLOKFLA
RKMILGILQILFIMPEFVILFKLPPPEFVLEFPGFVMSGLINSGLAFL
IAVRKTEETLIIISRIINMFLSALGAINGILLFEGFLIDNYICGYSHNSCKAVT
VLFGIITLMTFSLIEFLISLPSILGCHSEDDCECC"

BASE COUNT 195 a 142 c 117 g 237 t
ORIGIN

Query Match 98.5%; Score 665.6; DB 9; Length 691;
Best Local Similarity 99.4%; Pred. No. 2.5e-145;
Matches 668; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tcaatccttcaaatatcacagacacacatgatgatacaagcagacagtcagg 60
DB 20 TCATCTCCTTCAAAATATACCGACACCATGATGATCAAGCAGCAGACAGTCGG 79
QY 61 tgttctggtatttctccagaatacactgcttcagaataagagtcacagaacttcag 120
DB 80 TGTTCCTGATATTCCTCCAGAAATCACGCTTCAGAAATATGATCCAGAACTTCAG 139
QY 121 coacagactttcaactcaagcccttcgcaaaaattatttctagaanaatgaaaatct 180
DB 140 CCAGACCTTTTCAACTCAAGCCCTTCGCAAAAATATTTCTGAGAAAATGAAATCT 199
QY 181 taggaataccagatccctglttgaatatagaccttcttcttgagttacttcctt 240
DB 200 TAGGAGCTATCCAGATCCGTTTGAATATGACCTTCTTTCGAGTATCTTCCTT 259
QY 241 tcaacttgtaaaacacatatacaagttcccttataatttcttcagagatccattct 300
DB 260 TCACCTTTTAAACCATATCCAGAGTTCCCTTATATTTCTTCAGATATCATTCCT 319
QY 301 ggggctcgtttgttcaatlaattcttgagcccttcctaattgcaatgaaagaaaacca 360
DB 320 GGGGCTCTGTTTGTCTCAATTAATCTGAGGCTTCCTTAATTCAGATGAAAACCA 379
QY 361 cagaacacctgaataatttgagcgaataaagaatcttcttagtgcctgagagaatag 420
DB 380 CAGAAACCTGATATATATGAGCCGAAATATGAAATTTCTTAGTCCCTGGAGCAATAG 439
QY 421 ctggaatacttcctcacaatttggttcaatctagatcaaacacacatttggttatt 480
DB 440 CTGGAATATTTCTCTCACATTTTGTTTCAATCTAGATCAAAACATCATTTTGTTAT 499
QY 481 ctcaacaaaatagtcagtgtaagcgttactgcttcttcttgagaaatttgatatac 540
DB 500 CTCACCAAAATAGTCAGTGTAGGCTGTACTGCTGTTCTTGGAATTTTGATTAAT 559
QY 541 tgaactcttcaagcattatgaataatcaattctctctccttctcaatttgagggtgcc 600
DB 560 TGATATCTTTCAGCATATTTGAAATATTCATTTCTGCTTCATATTTTGGGGGTGC 619
QY 601 actcagaagatttgatgttgaacaaattgttgaactagcacttgagaataaagattg 660
DB 620 ACTCAGAGATTGTGATGTGAACAATGTTGTAAGTACAGCTGAGAAATGAAGATG 679
QY 661 ttaaaaataaaa 672
DB 680 TTAATAATATAA 691

RESULT 5
AX179781 603 bp DNA PAT 06-AUG-2001
LOCUS AX179781 Sequence 1 from Patent W00146417.
DEFINITION AX179781
ACCESSION AX179781
VERSION AX179781.1 GI:15132144
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
1 (bases 1 to 603)
AUTHORS Walke, D.W. and Turner, C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL having homology to cd20 proteins and ige receptors
Patent: WO 0146417-A 1 28-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
source 1..603
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 162 a 125 c 104 g 212 t
ORIGIN

Query Match 89.2%; Score 603; DB 6; Length 603;
Best Local Similarity 100.0%; Pred. No. 1e-130;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 atggaatcaagacacagacacagtcggttcttgtaatttctccagaatacactgct 92
DB 1 ATGGAATTCAGACCGCACAGTCGGTGTTCGTGATTTCTCCAGAAATCACTGCT 60
QY 93 tgaataatgagtcacagaaacttccagccaagaccttcaactcaaaagcccttgcaa 152
DB 61 TCAGATATGAGTCCACAGAACTTTCAGCCAGACCTTTCAACATCAAGCCCTTGCAA 120
QY 153 aaattattgtcagaanaaatgaaaatccttaagagatataccagactcttggaattatg 212
DB 121 AAATATTTGTCAGAAAATGAAAATCTTAGGACTATACAGATCTCTGTTGGAATATG 180
QY 213 accttctcttggagttacttctcttccacttgtaaaacacatccaagtttccc 272
DB 181 ACCCTTTCTTGGAGTATCTTCCTTTACCTTGTTAAACCATATCCAGATGTTCC 240
QY 273 ttatattcttccagagatccattctggagctctggttcttcaattatcttgagcc 332
DB 241 TTTATATTTCTTCAGGATATCATTTGAGGCTCTGTTTGTCAATTAATCTGAGGC 300
QY 333 ttcctaattgcagtgaaaagaaaacacagaaactcagataataattgagccgaataag 392
DB 301 TTCCTAATTCAGTGAAGAAAACCAAGAAACCTGATATATTTGAGCCGAATATG 360
QY 393 aatctcttagtgcccttgagagaataagctggaatcaattctctcaatttggttccatc 452
DB 361 AATCTTTCTTAGTGCCCTGAGCAATAGCTGGAATCATTCCTCAATTTGGTTCAATC 420
QY 453 ctgagatcaaaactacatttggttacttctcaacaaaatagtcagtgtaagcgttact 512
DB 421 CTAGATCAAAACTACATTTTGTTATTTCTCACCAAAATAGTCAAGTGAAGCTGTTACT 480
QY 513 gtccgttcttggaaatttgatatacatgataagacttcaagcattatgaaattatcatt 572
DB 481 GTCCTGTTCTTGGGAATTTTGATTTGATTTGATGATGATTTGATTTGATTTGATTT 540
QY 573 tctctgccttctcaaatlttggtgtgccaactcagaagatttgatltggaacaattgtgt 632
DB 541 TCTCTGCCCTTCTCAATTTTGGGTGCGACCTCAGAGGATTTGATTTGTAACAATGTTGT 600
QY 633 tga 635
DB 601 TGA 603

RESULT 6
AX179787 450 bp DNA PAT 06-AUG-2001
LOCUS AX179787 Sequence 7 from Patent W00146417.
DEFINITION AX179787
ACCESSION AX179787
VERSION AX179787.1 GI:15132147
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 450)
AUTHORS Walke, D.W. and Turner, C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0146417-A 7 28-JUN-2001;
Lexicon Genetics Incorporated (US)

FEATURES
source location/Qualifiers

1..450
/organism="Homo sapiens"

BASE COUNT 120 a 94 c 77 g 159 t
ORIGIN

Query Match 51.4%; Score 347.4; DB 6; Length 450;
Best Local Similarity 97.0%; Pred. No. 4e-71;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

33 atggatcgaagcagccagacagtcggtgttctgtatcttccccaagaatcagct 92
1 ATGGATTCAAGACCGCACAGTCGGGTTCGTGTGTAATTCCTCCAGAAATCACTGCT 60
QY 93 tcagaatagatgcacagaacttcagccagacacttcaactcaagcccttgcaa 152
DB 61 TCAGATATGATGACCCACAACTTTCAGCCAGACCTTTCAACTCAAGCCCTTGCAA 120
QY 153 aaattatgtcagaaaataaactcctaggaactcagaatcctgttggaaatg 212
DB 121 AAATATTTGCTAGAAAATGAAAATCTTAGGACATATCCAGATCCTGTTGGAATTAG 180
QY 213 accttctcttgaggatctatcctcttcaactgttaaacatcagaagtttccc 272
DB 181 ACCTTCTCTTGGAGTATATCTTCCCTTACCTGTGTAAACCATATCCAAAGTTCCC 240
QY 273 ttataatctcttcaggatatacattctggtgctctgtttgtcaataatctgagcc 332
DB 241 TTTATATTTCTTTCAGGATATTCATTCGTGGGCTGTGTTTGTCAATTAATCTGGAGCC 300
QY 333 ttccataatgcagtgaaagaaaacacagaaactctgataataatggccgaataag 392
DB 301 TTCCTAATTGACGTGAAAAGAAAACACAGAACTCTGGAATTTGTATCAATTGATG 360
QY 393 aactc 397
DB 361 ACTTT 365

RESULT 7

LOCUS AX179785 417 bp DNA PAT 06-AUG-2001
DEFINITION Sequence 5 from Patent WO0146417.
ACCESSION AX179785
VERSION AX179785.1 GI:15132146
KEYWORDS
SOURCE
ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 417)
AUTHORS Walke, D.W. and Turner, C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0146417-A 5 28-JUN-2001;
Lexicon Genetics Incorporated (US)

FEATURES
source location/Qualifiers

1..417
/organism="Homo sapiens"

BASE COUNT 115 a 91 c 69 g 142 t
ORIGIN

Query Match 50.6%; Score 341.8; DB 6; Length 417;
Best Local Similarity 99.4%; Pred. No. 1e-70;
Matches 343; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 33 atggatcgaagcagccagacagtcggtgttctgtatcttccccaagaatcagct 92
DB 1 ATGGATTCAAGACCGCACAGTCGGGTTCGTGTGTAATTCCTCCAGAAATCACTGCT 60
QY 93 tcagaatagatgcacagaacttcagccagacacttcaactcaagcccttgcaa 152
DB 61 TCAGATATGATGACCCACAACTTTCAGCCAGACCTTTCAACTCAAGCCCTTGCAA 120
QY 153 aaattatgtcagaaaataaactcctaggaactcagaatcctgttggaaatg 212
DB 121 AAATATTTGCTAGAAAATGAAAATCTTAGGACATATCCAGATCCTGTTGGAATTAG 180
QY 213 accttctcttgaggatctatcctcttcaactgttaaacatcagaagtttccc 272
DB 181 ACCTTCTCTTGGAGTATATCTTCCCTTACCTGTGTAAACCATATCCAAAGTTCCC 240
QY 273 ttataatctcttcaggatatacattctggtgctctgtttgtcaataatctgagcc 332
DB 241 TTTATATTTCTTTCAGGATATTCATTCGTGGGCTGTGTTTGTCAATTAATCTGGAGCC 300
QY 333 ttccataatgcagtgaaagaaaacacagaaactctgataataatggccgaataag 377
DB 301 TTCCTAATTGACGTGAAAAGAAAACACAGAACTCTGATCAAA 345

RESULT 8

LOCUS AX179783 246 bp DNA PAT 06-AUG-2001
DEFINITION Sequence 3 from Patent WO0146417.
ACCESSION AX179783
VERSION AX179783.1 GI:15132145
KEYWORDS
SOURCE
ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 246)
AUTHORS Walke, D.W. and Turner, C.A.
TITLE Human membrane proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0146417-A 3 28-JUN-2001;
Lexicon Genetics Incorporated (US)

FEATURES
source location/Qualifiers

1..246
/organism="Homo sapiens"

BASE COUNT 60 a 45 c 47 g 94 t
ORIGIN

Query Match 36.4%; Score 246; DB 6; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.7e-47;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 390 atgaatctcttagtgccttgagagaacttggaatcattctccacattgtttc 449
DB 1 ATGAATCTTCTTAGTGCCCTGAGACCAATGCTGGAATCATCTCCTCAATTTGGTTTC 60
QY 450 atccatagatcaaaactaacttgggtgtatcttcaccaaaatgacgtgtaagctgt 509
DB 61 ATCCATGATCAAAACTACATCTTGTGTGTTATCTCCACAAAATGTCAGTGAAGCTGTT 120
QY 510 actgctcgttcttggaatttgattatcagatgacttcaagatttaaatattc 569
DB 121 ACTGTCGCTTCTTGGGAATTTGATTAATGATGAGACTTTCAGCATTAATTAATTTTC 180
QY 570 attctctgcttctcaatttgggtgacactcagagagatttgattgtaacaatgt 629
|||||

DB 181 ATTCTCTGCTTTCATTTTGGGGTCCACTGACAGCATTTGATTGTGAACATGCT 240
 QY 630 tgttga 635
 |||||
 Db 241 TGTGTA 246

RESULT 9
 AC027787
 LOCUS AC027787
 DEFINITION Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS
 SEQUENCE SAMPLING:
 AC027787
 AC027787.2 GI:9845160
 HTG: HTGS_PHASE0.
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 138097)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 15, clone RP11-196E16
 Unpublished
 2 (bases 1 to 138097)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
 Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G.,
 Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collumore,A., Cooke,P., DeArrelano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domingo,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., Lacombe,K., Lamazares,R., Landers,T., Lehoczy,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
 Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenda,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 18, 2000 this sequence version replaced gi:7382629.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L9008
 Center clone name: 196_E_16

NOTE: This record contains 163 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 663: contig of 663 bp in length

664 763: gap of 100 bp
 764 1464: contig of 701 bp in length
 1465 1564: gap of 100 bp
 1565 2268: contig of 704 bp in length
 2269 2368: gap of 100 bp
 2369 3042: contig of 674 bp in length
 3043 3142: gap of 100 bp
 3143 3820: contig of 678 bp in length
 3821 3920: gap of 100 bp
 3921 4626: contig of 706 bp in length
 4627 4726: gap of 100 bp
 4727 5423: contig of 697 bp in length
 5424 5523: gap of 100 bp
 5524 6234: contig of 711 bp in length
 6235 6334: gap of 100 bp
 6335 7040: contig of 706 bp in length
 7041 7140: gap of 100 bp
 7141 7837: contig of 697 bp in length
 7838 7937: gap of 100 bp
 7938 8632: contig of 695 bp in length
 8633 8732: gap of 100 bp
 8733 9430: contig of 698 bp in length
 9431 9530: gap of 100 bp
 9531 10214: contig of 684 bp in length
 10215 10314: gap of 100 bp
 10315 10980: contig of 666 bp in length
 10981 11080: gap of 100 bp
 11081 11783: contig of 703 bp in length
 11784 11883: gap of 100 bp
 11884 12582: contig of 699 bp in length
 12583 12682: gap of 100 bp
 12683 13389: contig of 707 bp in length
 13390 13489: gap of 100 bp
 13490 14186: contig of 697 bp in length
 14187 14286: gap of 100 bp
 14287 14996: contig of 710 bp in length
 14997 15096: gap of 100 bp
 15097 15796: contig of 700 bp in length
 15797 15896: gap of 100 bp
 15897 16583: contig of 687 bp in length
 16584 16683: gap of 100 bp
 16684 17376: contig of 683 bp in length
 17377 17476: gap of 100 bp
 17477 18178: contig of 702 bp in length
 18179 18278: gap of 100 bp
 18279 18962: contig of 684 bp in length
 18963 19062: gap of 100 bp
 19063 19673: contig of 611 bp in length
 19674 19773: gap of 100 bp
 19774 20436: contig of 663 bp in length
 20437 20536: gap of 100 bp
 20537 21233: contig of 697 bp in length
 21234 21333: gap of 100 bp
 21334 22042: contig of 709 bp in length
 22043 22142: gap of 100 bp
 22143 22808: contig of 666 bp in length
 22809 22908: gap of 100 bp
 22909 23595: contig of 687 bp in length
 23596 23695: gap of 100 bp
 23696 24391: contig of 666 bp in length
 24392 24491: gap of 100 bp
 24492 25190: contig of 699 bp in length
 25191 25290: gap of 100 bp
 25291 25981: contig of 691 bp in length
 25982 26081: gap of 100 bp
 26082 26765: contig of 684 bp in length
 26766 26865: gap of 100 bp
 26866 27563: contig of 698 bp in length
 27564 27663: gap of 100 bp
 27664 28377: contig of 714 bp in length
 28378 28477: gap of 100 bp
 28478 29184: contig of 707 bp in length
 29185 29284: gap of 100 bp

29285	29983:	contlg of 699 bp	in length
29984	30083:	gap of 100 bp	
30084	30775:	contlg of 692 bp	in length
30786	30875:	gap of 100 bp	
30876	31565:	contlg of 690 bp	in length
31566	31665:	gap of 100 bp	
31666	32360:	contlg of 695 bp	in length
32361	32460:	gap of 100 bp	
32461	33157:	contlg of 697 bp	in length
33158	33257:	gap of 100 bp	
33258	33951:	contlg of 694 bp	in length
33952	34051:	gap of 100 bp	
34052	34735:	contlg of 684 bp	in length
34735	34835:	gap of 100 bp	
34836	35483:	contlg of 648 bp	in length
35484	35583:	gap of 100 bp	
35584	36308:	contlg of 725 bp	in length
36309	36408:	gap of 100 bp	
36409	37108:	contlg of 700 bp	in length
37109	37208:	gap of 100 bp	
37209	37892:	contlg of 684 bp	in length
37893	37992:	gap of 100 bp	
37993	38672:	contlg of 680 bp	in length
38673	38772:	gap of 100 bp	
38773	39475:	contlg of 703 bp	in length
39476	39575:	gap of 100 bp	
39576	40290:	contlg of 715 bp	in length
40291	40390:	gap of 100 bp	
40391	41087:	contlg of 697 bp	in length
41088	41187:	gap of 100 bp	
41188	41891:	contlg of 704 bp	in length
41892	41991:	gap of 100 bp	
41992	42679:	contlg of 688 bp	in length
42680	42779:	gap of 100 bp	
42780	43500:	contlg of 721 bp	in length
43501	43600:	gap of 100 bp	
43601	44295:	contlg of 695 bp	in length
44296	44395:	gap of 100 bp	
44396	45088:	contlg of 693 bp	in length
45089	45188:	gap of 100 bp	
45189	45880:	contlg of 692 bp	in length
45881	45980:	gap of 100 bp	
45981	46683:	contlg of 703 bp	in length
46684	46783:	gap of 100 bp	
46784	47481:	contlg of 698 bp	in length
47482	47581:	gap of 100 bp	
47582	48290:	contlg of 709 bp	in length
48291	48390:	gap of 100 bp	
48391	49077:	contlg of 687 bp	in length
49078	49177:	gap of 100 bp	
49178	49879:	contlg of 702 bp	in length
49880	49979:	gap of 100 bp	
49980	50684:	contlg of 705 bp	in length
50685	50784:	gap of 100 bp	
50785	51507:	contlg of 723 bp	in length
51508	51607:	gap of 100 bp	
51608	52320:	contlg of 713 bp	in length
52321	52420:	gap of 100 bp	
52421	53133:	contlg of 713 bp	in length
53134	53233:	gap of 100 bp	
53234	53933:	contlg of 700 bp	in length
53934	54033:	gap of 100 bp	
54034	54703:	contlg of 670 bp	in length

Query Match	27.48;	Score 185.4;	DB 2;	Length 138097;
Best Local Similarity	96.98;	Pred. No. 3.8e-33;		
Matches 189; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;

[illegible]

Df 109624 TGTTCCTGCGTATTTCCCTCCAAATACACTGGTTAGAAATAGATCCACAGAACTTCAG 1096633
Oy 121 ccacgacccttctaactcaaaagccctcgcataaatattatcgtagaanaaaygaactc 180
Df 109684 CCACGACCTTTTCAACATCAAAAGCCCCCTGCATAAATTATTTGGTCAGAAAAATGAANAATCT 1097434
Oy 181 taaggacctaccaga 195
Df 109744 TAGGGGTAAAGTAAGA 109758

RESULT	10
AC027787/c	
LOCUS	AC027787 138097 bp DNA HTG 18-AUG-2000
DEFINITION	Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS SEQUENCE SAMPLING.
ACCESSION	AC027787
VERSION	AC027787.2 GI:9845160
KEYWORDS	HTG; HTGS_PHASED.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	1 (bases 1 to 138097)
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE	Homo sapiens chromosome 15, clone RP11-166E16
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 138097)
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.

UNPUBLISHED
REFERENCES

2 (pages 1 to 138097)

Altman, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F., Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G., Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D., Gagegan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hages, B., Healdorf, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Marguis, N., McCarthy, M., McEwan, P., McGuck, A., McKernan, K., McNeethers, R., Meldrum, J., Menues, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olyar, T.M., Oliver, J., Peterson, K., Pierre, N., Pitsani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testeys, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

JOURNAL

COMMENT

Submitted (01-PR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 18, 2000 this sequence version replaced g1:7382629.

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L5008
Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 663: contig of 663 bp in length

664 763: gap of 100 bp
764 1464: contig of 701 bp in length
1465 1564: gap of 100 bp
1565 2268: contig of 704 bp in length
2269 2368: gap of 100 bp
2369 3042: contig of 674 bp in length
3043 3142: gap of 100 bp
3143 3820: contig of 678 bp in length
3821 3920: gap of 100 bp
3921 4626: contig of 706 bp in length
4627 4726: gap of 100 bp
4727 5423: contig of 697 bp in length
5424 5523: gap of 100 bp
5524 6234: contig of 711 bp in length
6235 6334: gap of 100 bp
6335 7040: contig of 706 bp in length
7041 7140: gap of 100 bp
7141 7837: contig of 697 bp in length
7838 7937: gap of 100 bp
7938 8632: contig of 695 bp in length
8633 8732: gap of 100 bp
8733 9430: contig of 698 bp in length
9431 9530: gap of 100 bp
9531 10214: contig of 684 bp in length
10215 10314: gap of 100 bp
10315 10980: contig of 666 bp in length
10981 11080: gap of 100 bp
11081 11783: contig of 703 bp in length
11784 11883: gap of 100 bp
11884 12582: contig of 699 bp in length
12583 12682: gap of 100 bp
12683 13389: contig of 707 bp in length
13390 13489: gap of 100 bp
13490 14186: contig of 697 bp in length
14187 14286: gap of 100 bp
14287 14996: contig of 710 bp in length
14997 15096: gap of 100 bp
15097 15796: contig of 700 bp in length
15797 15896: gap of 100 bp
15897 16583: contig of 687 bp in length
16584 16683: gap of 100 bp
16684 17376: contig of 693 bp in length
17377 17476: gap of 100 bp
17477 18178: contig of 702 bp in length
18179 18278: gap of 100 bp
18279 18962: contig of 684 bp in length
18963 19062: gap of 100 bp
19063 19673: contig of 611 bp in length
19674 19773: gap of 100 bp
19774 20436: contig of 663 bp in length
20437 20536: gap of 100 bp
20537 21333: contig of 697 bp in length
21334 21333: gap of 100 bp
21334 22042: contig of 709 bp in length
22043 22142: gap of 100 bp
22143 22808: contig of 666 bp in length
22809 22908: gap of 100 bp
22909 23595: contig of 687 bp in length
23596 23695: gap of 100 bp
23696 24391: contig of 696 bp in length
24392 24491: gap of 100 bp
24492 25190: contig of 699 bp in length
25191 25290: gap of 100 bp
25291 25981: contig of 691 bp in length
25982 26081: gap of 100 bp
26082 26765: contig of 684 bp in length
26766 26865: gap of 100 bp
26866 27563: contig of 698 bp in length

27564 27663: gap of 100 bp
27664 28377: contig of 714 bp in length
28378 28477: gap of 100 bp
28478 29184: contig of 707 bp in length
29185 29284: gap of 100 bp
29285 29983: contig of 699 bp in length
29984 30083: gap of 100 bp
30084 30775: contig of 692 bp in length
30776 30875: gap of 100 bp
30876 31565: contig of 690 bp in length
31566 31665: gap of 100 bp
31666 32360: contig of 695 bp in length
32361 32460: gap of 100 bp
32461 33157: contig of 697 bp in length
33158 33257: gap of 100 bp
33258 33951: contig of 694 bp in length
33952 34051: gap of 100 bp
34052 34735: contig of 684 bp in length
34736 34835: gap of 100 bp
34836 35483: contig of 648 bp in length
35484 35583: gap of 100 bp
35584 36308: contig of 725 bp in length
36309 36408: gap of 100 bp
36409 37108: contig of 700 bp in length
37109 37208: gap of 100 bp
37209 37892: contig of 684 bp in length
37893 37992: gap of 100 bp
37993 38672: contig of 680 bp in length
38673 38772: gap of 100 bp
38773 39475: contig of 703 bp in length
39476 39575: gap of 100 bp
39576 40290: contig of 715 bp in length
40291 40390: gap of 100 bp
40391 41087: contig of 697 bp in length
41088 41187: gap of 100 bp
41188 41891: contig of 704 bp in length
41892 41991: gap of 100 bp
41992 42679: contig of 688 bp in length
42680 42779: gap of 100 bp
42780 43500: contig of 721 bp in length
43501 43600: gap of 100 bp
43601 44295: contig of 695 bp in length
44296 44395: gap of 100 bp
44396 45088: contig of 693 bp in length
45089 45188: gap of 100 bp
45189 45880: contig of 692 bp in length
45881 45980: gap of 100 bp
45981 46683: contig of 703 bp in length
46684 46783: gap of 100 bp
46784 47481: contig of 698 bp in length
47482 47581: gap of 100 bp
47582 48290: contig of 709 bp in length
48291 48390: gap of 100 bp
48391 49077: contig of 687 bp in length
49078 49177: gap of 100 bp
49178 49879: contig of 702 bp in length
49880 49979: gap of 100 bp
49980 50684: contig of 705 bp in length
50685 50784: gap of 100 bp
50785 51507: contig of 723 bp in length
51508 51607: gap of 100 bp
51608 52320: contig of 713 bp in length
52321 52420: gap of 100 bp
52421 53133: contig of 713 bp in length
53134 53233: gap of 100 bp
53234 53933: contig of 700 bp in length
53934 54033: gap of 100 bp
54034 54703: contig of 670 bp in length

Query Match 27.4%; Score 185.4; DB 2; Length 138097;
Best Local Similarity 96.9%; Pred. No. 3.8e-33;
Matches 189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;


```

QY 1 tcatctccttcaaatatcaccgacacatcatgattcagaacgacacagtcgag 60
    |||||||
DB 112577 TCATCTCCTTCAAAATTATACACGACATCATGATTCAGACACCGACACAGTCGCG 112518
QY 61 tttcttggtatltccctccagaatactgcttcagaataatgctccacagaacttcag 120
    |||||||
DB 112517 TCTTTCTGTAATTTCTCCAGAAATCAGCTTCAGAAATATGATGTCACAGAACTTTCAG 112458
QY 121 ccagcacccttcaactcaagcccttcgaaataatttctgtagaanaatgaanaatct 180
    |||||||
DB 112457 CCACGACCTTTTCAACTCAAGCCCTTGCAGAAATATTTCTGTAGAAAATGAAATCT 112398
QY 181 taggagactaccaga 195
    |||||
DB 112397 TAGGGGTAGTAAGA 112383

RESULT 11
AC015840 147788 bp DNA HTG 27-APR-2000
LOCUS Homo sapiens clone RP11-24D1, WORKING DRAFT SEQUENCE, 5 unordered
DEFINITION pieces
AC015840
VERSION AC015840.2 GI:7657730
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 147788)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens, clone RP11-24D1
TITLE Unpublished
AUTHORS 2 (bases 1 to 147788)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckert,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearliano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gairdya,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,N.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McDonald,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Sudmanian,A., Talamas,J.,
Testa,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced g1:6446828.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: 14434
Center clone name: 24.D.1
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 136347 bases at least Q40
Consensus quality: 143600 bases at least Q30
Consensus quality: 146212 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 147388; sum-of-contigs
Quality coverage: 5.2 in Q20 bases; agarose-fp

```

```

Quality coverage: 7.1 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 6594: contig of 6594 bp in length
* 6595 6694: gap of 100 bp
* 6695 16187: contig of 9493 bp in length
* 16188 16287: gap of 100 bp
* 16288 25089: contig of 8802 bp in length
* 25090 25189: gap of 100 bp
* 25190 63923: contig of 38734 bp in length
* 63924 64023: gap of 100 bp
* 64024 147788: contig of 83765 bp in length.
FEATURES
Source
1..147788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-24D1"
/clone_1b="RPCT-11 Human Male BAC"
1..6594
/note="assembly-fragment
clone_end:17
vector_side:left"
6595..16187
/note="assembly-fragment"
16288..25089
/note="assembly-fragment
clone_end:SP6
vector_side:right"
25190..63923
/note="assembly-fragment"
64024..147788
/note="assembly-fragment"
BASE COUNT 45466 a 29482 c 28861 g 43569 t 410 others
ORIGIN

Query Match 27.4%; Score 185.4; DB 2; Length 147788;
Best Local Similarity 96.9%; Pred. No. 3.8e-33;
Matches 189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 tcatctccttcaaatatcaccgacacatcatgattcagaacgacacagtcgag 60
    |||||||
DB 128015 TCATCTCCTTCAAAATTATACACGACATCATGATTCAGACACCGACACAGTCGCG 128074
QY 61 tttcttggtatltccctccagaatactgcttcagaataatgctccacagaacttcag 120
    |||||||
DB 128075 TCTTTCTGTAATTTCTCCAGAAATCAGCTTCAGAAATATGATGTCACAGAACTTTCAG 128134
QY 121 ccagcacccttcaactcaagcccttcgaaataatttctgtagaanaatgaanaatct 180
    |||||||
DB 128135 CCACGACCTTTTCAACTCAAGCCCTTGCAGAAATATTTCTGTAGAAAATGAAATCT 128194
QY 181 taggagactaccaga 195
    |||||
DB 128195 TAGGGGTAGTAAGA 128209

RESULT 12
AP003127 161039 bp DNA HTG 30-JAN-2001
LOCUS Homo sapiens chromosome 11 clone RP11-7106 map 11q, WORKING DRAFT
DEFINITION SEQUENCE, 15 unordered pieces.
ACCESSION AP003127
VERSION AP003127.1 GI:12597183
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-7106.

```



```

OY      121  ccagcactttcaactcaaaagcccttcgcaaaatatttgcctagaaaaatgaaatc 180
OY      181  taggagactaccaga 195
DB 149530  CCACGACCTTTTCACACTCAAAAGCCCTTGCAAAATATTCTGCTAGAAAAATGAAATCT 149471
OY      181  taggagactaccaga 195
DB 149470  TAGGGCTAAGTAAAGA 149456

RESULT  13
AP001034 166804 bp DNA HTG 14-NOV-2000
LOCUS     Homo sapiens chromosome 11 clone RP11-729B4 map 11q12, WORKING
DEFINITION
DRAFT SEQUENCE, 19 unordered pieces.
ACCESSION AP001034.4 GI:11176992
VERSION    HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS   Homo sapiens DNA, clone:RP11-729B4.
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 166804)
            Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Hattori,M., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
            Homo sapiens 166,804 genomic DNA of 11q12
            Published Only in Database (2000) in press
            2 (bases 1 to 166804)
            Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
            Direct Submission
            Submitted (06-JAN-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hnp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            On Nov 14, 2000 this sequence version replaced gi:8117704.

COMMENT   ----- Genome Center
            Center: RIKEN Genomic Sciences Center(GSC)
            Web site: http://hnp.gsc.riken.go.jp/
            Contact: hattori@gsc.riken.go.jp
            ----- Project Information
            Center project name: Humdraft11
            Center clone name: RP11-729B4
            ----- Summary Statistics
            Sequencing vector: PCR products; 100% of reads
            Chemistry: Dye-terminator ET-amersham; 100% of reads
            Assembly program: Phrap; version 0.990329
            Consensus quality: 162254 bases at least Q40
            Consensus quality: 163858 bases at least Q30
            Consensus quality: 164544 bases at least Q20
            Insert size: 165004; sum-of-contents
            Quality coverage: 9.28x in Q20 bases; sum-of-contents

NOTE: This is a 'working draft' sequence. It currently consists of
19 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.

```

```

144754 149354 contig of 4601 bp in length
149455 154522 contig of 5068 bp in length
154623 155919 contig of 1297 bp in length
156020 159504 contig of 3485 bp in length
159605 159679 contig of 75 bp in length
159780 162418 contig of 2639 bp in length
162519 165383 contig of 2865 bp in length
165484 166804 contig of 1321 bp in length.

NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
22398: contig of 22398 bp in length
22399 22498: gap of 100 bp
22499 45238: contig of 22740 bp in length
45239 45338: gap of 100 bp
45339 64278: contig of 18940 bp in length
64279 64378: gap of 100 bp
64379 77579: contig of 13201 bp in length
77580 77679: gap of 100 bp
77680 93718: contig of 16039 bp in length
93719 93818: gap of 100 bp
93819 108057: contig of 14239 bp in length
108058 108157: gap of 100 bp
108158 117006: contig of 8849 bp in length
117007 117106: gap of 100 bp
117107 126302: contig of 9196 bp in length
126303 126402: gap of 100 bp
126403 133479: contig of 7077 bp in length
133480 133579: gap of 100 bp
133580 140425: contig of 6646 bp in length
140426 140525: gap of 100 bp
140526 144653: contig of 4128 bp in length
144654 144753: gap of 100 bp
144754 149354: contig of 4601 bp in length
149355 149454: gap of 100 bp
149455 154522: contig of 5068 bp in length
154523 154622: gap of 100 bp
154623 155919: contig of 1297 bp in length
155920 156019: gap of 100 bp
156020 159504: contig of 3485 bp in length
159505 159604: gap of 100 bp
159605 159679: contig of 75 bp in length
159680 159779: gap of 100 bp
159780 162418: contig of 2639 bp in length
162419 162518: gap of 100 bp
162519 165383: contig of 2865 bp in length
165384 165483: gap of 100 bp
165484 166804: contig of 1321 bp in length.

FEATURES
source
1. 166804
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="11"
   /map="11q12"
   /clone="RP11-729B4"
1. 22398
   /note="assembly-fragment"
22499. 45238
   /note="assembly-fragment"
45339. 64278
   /note="assembly-fragment"
64379. 77579
   /note="assembly-fragment"
77680. 93718
   /note="assembly-fragment"
93819. 108057
   /note="assembly-fragment"
108158. 117006

```


Best Local Similarity 96.9%; Pred. No. 3.9e-33;
Matches 189; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 tcatctcccttaaatctaccagacacatgatgattcaagacccgacagctccg 60
Db 144132 TCATCTCCCTTAATAATATACCGACCATCATGTGATTCAAGACCCGACACAGACTTCG 144073
QY 61 tcttcttgatctcctccagaatcactgcttcagatatagtcacagacttcag 120
Db 144072 TGTTCCTGATTTCTCTCCAGAAATCAGCTTCAGAAATATGATGCCAGAACTTTCAG 144013
QY 121 ccagacaccttccaactcaagcccttcgaaaattattgctagaaaatgaaatct 180
Db 144012 CCAGACACCTTTCAACTCAAGCCCTTCGCAAAAATTATTCTGTAGAAAATGAAATCT 143953
QY 181 taggactactcaga 195
Db 143952 TAGGGGTAGTAGA 143938

RESULT 15

LOCUS AP003127 161039 bp DNA HTG 30-JAN-2001
DEFINITION Homo sapiens chromosome 11 clone RP11-710G6 map 11q, WORKING DRAFT
ACCESSION AP003127
VERSION AP003127.1 GI:12597183
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-710G6.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 161039)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H., and Sakaki, Y.
Homo sapiens 161,039 genomic DNA of 11q
Published Only in Database (2001) In press
2 (bases 1 to 161039)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H., and Sakaki, Y.
Direct Submission

TITLE Journal
JOURNAL Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
AUTHORS
COMMENT

Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp

Project Information
Center project name: Humdraf11
Center clone name: RP11-710G6

Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 157456 bases at least Q40
Consensus quality: 158960 bases at least Q30
Consensus quality: 159457 bases at least Q20
Insert size: 159639; sum-of-ctgigs
Quality coverage: 8.61x in Q20 bases; sum-of-ctgigs

NOTE: This is a 'working draft' sequence. It currently consists of
15 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

1 42322 contig of 42322 bp in length
42423 63418 contig of 20996 bp in length

63519 77969 contig of 14451 bp in length
78070 93182 contig of 15113 bp in length
93283 104093 contig of 10811 bp in length
104194 115702 contig of 11509 bp in length
115803 125183 contig of 9381 bp in length
125284 132558 contig of 7275 bp in length
132659 137219 contig of 4561 bp in length
137320 143276 contig of 5546 bp in length
143377 148922 contig of 5137 bp in length
149023 154159 contig of 5137 bp in length
154260 157686 contig of 3427 bp in length
157787 159234 contig of 1448 bp in length
159335 161039 contig of 1705 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 42323 42422: gap of 100 bp
* 42423 63418: contig of 20996 bp in length
* 63419 63518: gap of 100 bp
* 63519 77969: contig of 14451 bp in length
* 77970 78069: gap of 100 bp
* 78070 93182: contig of 15113 bp in length
* 93183 93282: gap of 100 bp
* 93283 104093: contig of 10811 bp in length
* 104094 104193: gap of 100 bp
* 104194 115702: contig of 11509 bp in length
* 115703 115802: gap of 100 bp
* 115803 125183: contig of 9381 bp in length
* 125184 125283: gap of 100 bp
* 125284 132558: contig of 7275 bp in length
* 132559 132658: gap of 100 bp
* 132659 137219: contig of 4561 bp in length
* 137220 137319: gap of 100 bp
* 137320 143276: contig of 5546 bp in length
* 143277 143376: gap of 100 bp
* 143377 148922: contig of 5137 bp in length
* 148923 149022: gap of 100 bp
* 149023 154159: contig of 5137 bp in length
* 154160 154259: gap of 100 bp
* 154260 157686: contig of 3427 bp in length
* 157687 157786: gap of 100 bp
* 157787 159234: contig of 1448 bp in length
* 159235 159334: gap of 100 bp
* 159335 161039: contig of 1705 bp in length.
FEATURES
Source
1. 161039
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-710G6"
1. 42322
/note="assembly-fragment"
42423. 63418
/note="assembly-fragment"
63519. 77969
/note="assembly-fragment"
78070. 93182
/note="assembly-fragment"
93283. 104093
/note="assembly-fragment"
104194. 115702
/note="assembly-fragment"
115803. 125183
/note="assembly-fragment"
125284. 132558
/note="assembly-fragment"

misc_feature
1. 42322
/note="assembly-fragment"
42423. 63418
/note="assembly-fragment"
63519. 77969
/note="assembly-fragment"
78070. 93182
/note="assembly-fragment"
93283. 104093
/note="assembly-fragment"
104194. 115702
/note="assembly-fragment"
115803. 125183
/note="assembly-fragment"
125284. 132558
/note="assembly-fragment"

```

misc_feature      132659..137219
                  /note="assembly_fragment clone_end:sp6 vector_side:left"
misc_feature      137320..143276
                  /note="assembly_fragment"
misc_feature      143377..148922
                  /note="assembly_fragment clone_end:T7 vector_side:left"
misc_feature      149023..154159
                  /note="assembly_fragment"
misc_feature      154260..157686
                  /note="assembly_fragment"
misc_feature      157787..159234
                  /note="assembly_fragment"
misc_feature      159335..161039
                  /note="assembly_fragment"
BASE COUNT      49155 a 31166 c 31523 g 47795 t 1400 others
ORIGIN

```

```

Query Match      23.2%; Score 157; DB 2; Length 161039;
Best Local Similarity 91.7%; Pred. No. 1.6e-26;
Matches 166; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY      368      tctgataataggacgcaataaataatcctctcttagtgcctcgagagcaatagctggaat 427
          |||||||
Db      45928      TCAGATAATATTGACCCGATAATATTTCTTAGTGCCCTGGAGCAATAGCTGAAT 45987

OY      428      catctcctcacatttggttcatctctagatcaaaactacatttggttattctcacca 487
          |||||||
Db      45988      CATCTCCTCACATTGGTTTCATCTAGATCAAAACTACATTGTGGTATTCTCACCA 46047

OY      488      aaatagtcagtgtaaggctgcttactgctctgcttcttggaatttgattacattgac 547
          |||||||
Db      46048      AAATAGTCAGTGTAAGCTGTACTGCTCTGTCTTGGTAGTAGTTCATTATTAGAG 46107

OY      548      t 548
          |
Db      46108      T 46108

```

Search completed: November 28, 2001, 23:52:04
 Job time: 8367 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 00:23:54 ; Search time 204.6 Seconds
(without alignments)
2832,609 Million cell updates/sec

Title: US-09-735-712-9

Perfect score: 676
Sequence: 1 tcatctcccttcaattatc.....tjgttcaataataaaaaa 676

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 930621 seqs, 428662619 residues

tal number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	676	100.0	676	22	AA504283
2	668	98.8	747	22	AAH64741
3	603	89.2	603	22	AA504279
4	347.4	51.4	450	22	AA504282
5	341.8	50.6	417	22	AA504281
6	312.8	46.3	372	21	AA503080
7	246	36.4	246	22	AA504280
8	190.4	28.2	468	22	AAH64745
9	108.8	16.1	936	22	AA58252
10	108.8	16.1	936	22	AA58252
11	108.8	16.1	936	22	AA58254

c	12	108.8	16.1	936	22	AA58254	Oligonucleotide D1
c	13	108.8	16.1	936	22	AA58257	Oligonucleotide D1
c	14	108.8	16.1	936	22	AA58257	Oligonucleotide D1
c	15	108.8	16.1	936	22	AA58259	Oligonucleotide D2
c	16	108.8	16.1	936	22	AA58259	Oligonucleotide D2
c	17	108.8	16.1	936	22	AA58259	Oligonucleotide D2
c	18	108.8	16.1	936	22	AA58252	Oligonucleotide D2
c	19	108.8	16.1	936	22	AA58252	Oligonucleotide D2
c	20	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	21	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	22	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	23	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	24	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	25	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	26	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	27	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	28	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	29	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	30	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	31	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	32	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	33	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	34	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	35	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	36	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	37	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	38	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	39	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	40	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	41	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	42	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	43	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	44	108.8	16.1	936	22	AA58255	Oligonucleotide D1
c	45	108.8	16.1	936	22	AA58255	Oligonucleotide D1

ALIGNMENTS

RESULT	1
AA504283	standard; DNA; 676 BP.
ID	AA504283;
AC	AA504283;
DT	26-SEP-2001 (first entry)
XX	
DE	DNA sequence encoding novel human membrane protein.
XX	
KW	Human: membrane protein; membrane receptor; Ige receptor; CD20;
KW	physiological disorder; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200146417-A2.
XX	
PD	28-JUN-2001.
XX	
PF	12-DEC-2000; 2000WO-US33742.
XX	
PR	22-DEC-1999; 99US-0171567.
XX	
RA	(LEXI-) LEXICON GENETICS INC.
XX	
PI	Walke DW, Turner CA;
XX	
DR	WPI; 2001-408646/43.
XX	
PT	Polynucleotide encoding novel human membrane protein, useful for
PT	identifying agonist, antagonist or modifiers or for producing
PT	antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT	applications
XX	

PS Disclosure: Page 31-32; 32pp; English.

XX The present sequence encoding for a novel human membrane protein
CC includes the flanking 5'- and 3'- sequences. Four amino acid sequences
CC for novel human membrane proteins (AAU01210-AAU01213) are given in the
CC present invention. These membrane proteins share structural similarity
CC with membrane receptors such as the IGF receptor and mammalian CD20.
CC The novel human membrane proteins are useful for identifying agonist,
CC antagonist and modulators of the membrane proteins, and for producing
CC antibodies specific to the membrane proteins. The membrane proteins can
CC be used for diagnosis, drug screening, pharmacogenomic applications,
CC clinical trial monitoring and the treatment of physiological disorders
CC and diseases. The polynucleotides encoding the membrane proteins can be
CC used to generate PCR primers or probes to identify mutations associated
CC with a particular disease.
XX

Sequence 676 BP; 193 A; 140 C; 112 G; 231 T; 0 other;

Query Match 100.0%; Score 676; DB 22; Length 676;
Best Local Similarity 100.0%; Pred. No. 3,7e-174;
Matches 676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tcattcccttcaaatatccacgacacatcatgattcaagcaccgacacagtcg 60
DB 1 tcattcccttcaaatatccacgacacatcatgattcaagcaccgacacagtcg 60
QY 61 ttttctggtatttccctcagaatcactgtctcagaatgatgagtcacagacttcag 120
DB 61 ttttctggtatttccctcagaatcactgtctcagaatgatgagtcacagacttcag 120
QY 121 ccaagacctttcaactaaagcccttgcaaaaattttgtcgtagaataatgaatct 180
DB 121 ccaagacctttcaactaaagcccttgcaaaaattttgtcgtagaataatgaatct 180
QY 181 taggagacatccagactcgtttggaattatgacctttcttttggaattatcttcctt 240
DB 181 taggagacatccagactcgtttggaattatgacctttcttttggaattatcttcctt 240
QY 241 tcacctgtttaaaccataatcagaagttcccttataattcttccagatatacatct 300
DB 241 tcacctgtttaaaccataatcagaagttcccttataattcttccagatatacatct 300
QY 301 ggggctcgtttgttctatattcttgcagaccttcctatgctgagtgagaagaaacca 360
DB 301 ggggctcgtttgttctatattcttgcagaccttcctatgctgagtgagaagaaacca 360
QY 361 cagaacctgtataatattgagccgaataatgaatcttctagtgaccttgagcaatag 420
DB 361 cagaacctgtataatattgagccgaataatgaatcttctagtgaccttgagcaatag 420
QY 421 ctggaatattctcctcactatgttctcactcagatcaaaaactactgtgtatt 480
DB 421 ctggaatattctcctcactatgttctcactcagatcaaaaactactgtgtatt 480
QY 481 ctcaacaaatagctgagtgagctgttactgtcctgttttggaatttgattat 540
DB 481 ctcaacaaatagctgagtgagctgttactgtcctgttttggaatttgattat 540
QY 541 tgatgacttcagatattatgaattatcattctcctccttcacatttggtggtcc 600
DB 541 tgatgacttcagatattatgaattatcattctcctccttcacatttggtggtcc 600
QY 601 actcagaggtatgattgtgaaacaatgttctgactagcacttgagaataagatgtg 660
DB 601 actcagaggtatgattgtgaaacaatgttctgactagcacttgagaataagatgtg 660
QY 661 ttaataataaaaaaa 676
DB 661 ttaataataaaaaaa 676

RESULT 2

AAH64741
ID AAH64741 standard; cDNA; 747 BP.

XX AAH64741;

AC 11-SEP-2001 (first entry)

DE Human secreted protein cDNA, SEQ ID NO: 17.

DE Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;

KW GENSET, ss.

OS Homo sapiens.

PN WO200142451-A2.

PD 14-JUN-2001.

PF 07-DEC-2000; 2000WO-IB01938.

PR 08-DEC-1999; 99US-0169629.

PR 06-MAR-2000; 2000US-0187470.

XX (GENSET) GENSET.

XX Dumas Milne Edwards J, Bougueleret L, Jobert S;

XX WPI; 2001-367870/38.

DR P-PSDB; AAG89138.

PS Claim 7; Page 576-577; 921pp; English.

CC The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC GENSET gene expression by rectifying mutations or deletions in a
CC patient's genome that affect the activity of GENSET or by supplementing
CC the patient's own production of GENSET polypeptides. Conversely,
CC antisense nucleic acid molecules may be administered to down regulate
CC GENSET expression by binding with the cells' own genes and preventing
CC their expression. The sense and antisense nucleic acids may also be
CC used as DNA probes in diagnostic assays to detect and quantify the
CC presence of similar nucleic acid sequences in samples, and hence to
CC determine which patients may be in need of restorative therapy.
CC The GENSET polypeptides may be used as antigens in the production of
CC antibodies and in assays to identify modulators (agonists and
CC antagonists) of GENSET polypeptide expression and activity. The
CC present sequence is a GENSET nucleic acid of the invention.

Sequence 747 BP; 240 A; 145 C; 121 G; 241 T; 0 other;

Query Match 98.8%; Score 668; DB 22; Length 747;
Best Local Similarity 99.3%; Pred. No. 5,7e-172;
Matches 671; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tcattcccttcaaatatccacgacacatcatgattcaagcaccgacacagtcg 60
DB 1 tcattcccttcaaatatccacgacacatcatgattcaagcaccgacacagtcg 60
QY 61 ttttctggtatttccctcagaatcactgtctcagaatgatgagtcacagacttcag 120
DB 61 ttttctggtatttccctcagaatcactgtctcagaatgatgagtcacagacttcag 120
QY 101 ttttctggtatttccctcagaatcactgtctcagaatgatgagtcacagacttcag 160
DB 101 ttttctggtatttccctcagaatcactgtctcagaatgatgagtcacagacttcag 160
QY 121 ccaagacctttcaactaaagcccttgcaaaaattttgtcgtagaataatgaatct 180
DB 121 ccaagacctttcaactaaagcccttgcaaaaattttgtcgtagaataatgaatct 180
QY 161 ccaagacctttcaactaaagcccttgcaaaaattttgtcgtagaataatgaatct 220
DB 161 ccaagacctttcaactaaagcccttgcaaaaattttgtcgtagaataatgaatct 220


```
OY 181 taggacatccagatccctgttggaattatgacctttctttgttgagttatctctt 240
    |||||
Db 221 taggacatccagatccctgttggaattatgacctttctttgttgagttatctctt 280
OY 241 tcacctgttaaacacatcccaaggtcccttattatcttcttcaggatcatctc 300
    |||||
Db 281 tcacctgttaaacacatcccaaggtcccttattatcttcttcaggatcatctc 340
OY 301 ggggctctgtttgttcattatcttgagccttcttaattgcaaggaagaaca 360
    |||||
Db 341 ggggctctgtttgttcattatcttgagccttcttaattgcaaggaagaaca 400
OY 361 caaagacttgataattgagccgaataatgactctcttgagtcggagagcaatag 420
    |||||
Db 401 cagaactctgataattgagccgaataatgactctcttgagtcggagagcaatag 460
OY 421 ctggaatcatctctcacaattgtttcatcctagatcaaaactacattgtgtatt 480
    |||||
Db 461 ctggaatcatctctcacaattgtttcatcctagatcaaaactacattgtgtatt 520
OY 481 ctacccaataatagtcagtgtaagcgtctactgctctgttcttgggaatttgattacat 540
    |||||
Db 521 ctacccaataatagtcagtgtaagcgtctactgctctgttcttgggaatttgattacat 580
OY 541 ttagtacttcagcatattgattatattcttctgcttcttcaatttgggggtgc 600
    |||||
Db 581 ttagtacttcagcatattgattatattcttctgcttcttcaatttgggggtgc 640
OY 601 actcaagagattgtgattgtgaacaatgttgttgaactgaactgtgaagataaagattgtg 660
    |||||
Db 641 actcaagagattgtgattgtgaacaatgttgttgaactgaactgtgaagataaagattgtg 700
OY 661 ttaataataaaaaaa 676
    |||||
Db 701 ttaataataaaaaaa 716

RESULT 3
AAS04279
ID AAS04279 standard; cDNA; 603 BP.
XX
AC AAS04279;
DT 26-SEP-2001 (first entry)
DE Novel human membrane protein #1 cDNA sequence.
XX
Human; membrane protein; membrane receptor; Ige receptor; CD20;
physiological disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..603
FT FT /*tag= a
FT FT /product= "Human membrane protein #1"
FT FT /transl_except= (pos:595..600,aa:Cys)
XX
XX WO200146417-A2.
XX
XX 28-JUN-2001.
XX
XX 12-DEC-2000; 2000MO-US33742.
XX
XX 22-DEC-1999; 99US-0171567.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Walke DW, Turner CA;
XX
XX WPI: 2001-408646/43.
XX
XX P-PSDB; AA001210.
```

```
XX Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications
PS
XX
XX Claim 1; Page 29; 32pp; English.
CC
CC The present sequence encodes for novel human membrane protein #1.
CC Human membrane protein #1 is 1 of 4 human membrane proteins
CC (AA001210-AA001213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the Ige receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiological disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.
XX
XX Sequence 603 BP; 162 A; 125 C; 104 G; 212 T; 0 other;
SQ
Query Match 89.2%; Score 603; DB 22; Length 603;
Best Local Similarity 100.0%; Pred. No. 2.4e-154;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 33 atgattcaagcagccgacacagctccggtgttctgtattcttcctcaaatcactgtc 92
Db 1 atgattcaagcagccgacacagctccggtgttctgtattcttcctcaaatcactgtc 60
OY 93 tcagaataatgagtcacagaaacttccagccagcacttccaactcaaaagcccttgcaa 152
Db 61 tcagaataatgagtcacagaaacttccagccagcacttccaactcaaaagcccttgcaa 120
OY 153 aaattatgtcagaaataatgaataatccttaggagcattccagatccctgttgaattatg 212
Db 121 aaattatgtcagaaataatgaataatccttaggagcattccagatccctgttgaattatg 180
OY 213 acccttctcttgaggtatcttcccttccacttggttaaaacacatccaaagttccc 272
Db 181 acccttctcttgaggtatcttcccttccacttggttaaaacacatccaaagttccc 240
OY 273 ttataattcttcaagatatccattctggtgctctgttttgaattatcttgagcc 332
Db 241 ttataattcttcaagatatccattctggtgctctgttttgaattatcttgagcc 300
OY 333 ttcccaattgagtgaaagaaagaaacacagaaactctgataatttggccgaataatg 392
Db 301 ttcccaattgagtgaaagaaagaaacacagaaactctgataatttggccgaataatg 360
OY 393 aattctttagtgccctgagagcaatagctggaatcatctctccacatttggttcac 452
Db 361 aattctttagtgccctgagagcaatagctggaatcatctctccacatttggttcac 420
OY 453 ctatgacaaactacatttgggttaattctcaacaaatagtcagtgtaaggcttact 512
Db 421 ctatgacaaactacatttgggttaattctcaacaaatagtcagtgtaaggcttact 480
OY 513 gtccgtctctgggaatttgaattatgattgagacttcagatattggaattatctc 572
Db 481 gtccgtctctgggaatttgaattatgattgagacttcagatattggaattatctc 540
OY 573 tctctgcttctcaatttgggtgagcactcagaggaattggaattggaacaggtgtg 632
Db 541 tctctgcttctcaatttgggtgagcactcagaggaattggaattggaacaggtgtg 600
OY 633 tga 635
Db 601 tga 603
```


PA (LEXI-) LEXICON GENETICS INC.
 XX Walke DW, Turner CA;
 XX WPI: 2001-408646/43.
 DR P-PSDB: AAU01211.
 XX Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications
 XX
 PS Disclosure: Page 30; 32pp; English.
 XX
 CC The present sequence encodes for novel human membrane protein #2.
 CC Human membrane protein #2 is 1 of 4 human membrane proteins
 CC (AAU01210-AAU01213) given in the present invention. These membrane
 CC proteins share structural similarity with membrane receptors such as
 CC the Ige receptor and mammalian CD20. The novel human membrane proteins
 CC are useful for identifying agonists, antagonists and modulators of the
 CC membrane proteins, and for producing antibodies specific to the
 CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring
 CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.
 XX
 SQ Sequence 246 BP; 60 A; 45 C; 47 G; 94 T; 0 other;
 Query Match 36.4%; Score 246; DB 22; Length 246;
 Best Local Similarity 100.0%; Pred. No. 1.7e-57;
 Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 390 atgaattcttagtgagcctgagagcaatagtgatcattcctcattgtttc 449
 Db 1 atgaattcttagtgagcctgagagcaatagtgatcattcctcattgtttc 60
 QY 450 atcctagatcaaaaactacattgtgttatctccacaaaatagtcagtgtaagcgtt 509
 Db 61 atcctagatcaaaaactacattgtgttatctccacaaaatagtcagtgtaagcgtt 120
 QY 510 actgctgcgttcttgaggaaatttgatcattgactgacttcagcattatgaattatc 569
 Db 121 actgctgcgttcttgaggaaatttgatcattgactgacttcagcattatgaattatc 180
 QY 570 attctctgccttctcaatttgggtggtccactcagaggaattggtatgtaacaatgt 629
 Db 181 attctctgccttctcaatttgggtggtccactcagaggaattggtatgtaacaatgt 240
 QY 630 tgttga 635
 Db 241 tgttga 246
 RESULT 8
 AAH64745
 ID AAH64745 standard; cDNA; 468 BP.
 XX
 AC AAH64745;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Human secreted protein cDNA, SEQ ID NO: 21.
 XX
 KW Human: secreted protein; gene therapy; vaccine; treatment; diagnosis;
 KM GENSET; SS.
 XX Homo sapiens.
 XX OS
 XX PN WO200142451-A2.
 XX

PD 14-JUN-2001.
 XX
 XX 07-DEC-2000; 2000WO-IB01938.
 XX
 XX 08-DEC-1999; 99US-0169629.
 PR 06-MAR-2000; 2000US-0187470.
 XX
 XX (GEST) GENSET.
 PA Dumas Milne Edwards J, Bougueleret L, Jobert S;
 PI WPI: 2001-367870/38.
 DR P-PSDB: AAG89142.
 XX
 XX Full length GENSET human nucleic acids encoding potentially secreted
 PT proteins, useful in gene therapy and vaccination against a variety of
 PT diseases, and for diagnosis of those diseases -
 XX
 PS Claim 7; Page 581; 921pp; English.
 XX
 CC The invention relates to full length GENSET human nucleic acids encoding
 CC potentially secreted proteins. The nucleic acids and the polypeptides
 CC they encode may be used in the prevention, treatment and diagnosis of
 CC diseases associated with inappropriate GENSET gene expression. For
 CC example, they be used to treat disorders associated with decreased
 CC GENSET gene expression by rectifying mutations or deletions in a
 CC patient's genome that affect the activity of GENSET or by supplementing
 CC the patients own production of GENSET polypeptides. Conversely,
 CC antisense nucleic acid molecules may be administered to down regulate
 CC GENSET expression by binding with the cells' own genes and preventing
 CC their expression. The sense and antisense nucleic acids may also be
 CC used as DNA probes in diagnostic assays to detect and quantitate the
 CC presence of similar nucleic acid sequences in samples, and hence to
 CC determine which patients may be in need of restorative therapy.
 CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET nucleic acid of the invention.
 XX
 SQ Sequence 468 BP; 155 A; 94 C; 81 G; 138 T; 0 other;
 Query Match 28.2%; Score 190.4; DB 22; Length 468;
 Best Local Similarity 77.7%; Pred. No. 2.7e-42;
 Matches 230; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 1 tcaatccttcaaatatcaacgacacatcattgattcaagcagcagacagtcg 60
 Db 44 tcaatccttcaaatatcaacgacacatcattgattcaagcagcagacagtcg 103
 QY 61 tttctgttatcttcctcagaatacactgcttcagaatagatccacagaacttcag 120
 Db 104 tttctgttatcttcctcagaatacactgcttcagaatagatccacagaacttcag 163
 QY 121 ccaagacctttcaactaaagcccttgcaaaaataattgtcagaaaatgaaatct 180
 Db 164 ccaagacctttcaactaaagcccttgcaaaaataattgtcagaaaatgaaatct 223
 QY 181 taggagactcagagatcctgttggaattgactcttcttggtgggttacttcctt 240
 Db 224 taggagatccatctcgtggtctgtttgttcaattcaattcgtgagaccttcccaatgc 283
 QY 241 tcacctgttaaacacatatacgaagttcccttataattcttcagagatatca 296
 Db 284 agtgaagaagaaacacacagaactcttgggaatttggattatcatgtatgacttca 339
 RESULT 9
 AAF58252
 ID AAF58252 standard; DNA; 936 BP.
 XX
 AC AAF58252;
 XX

```

DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
XX gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000MO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
XX 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface.
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
XX acids each containing an electron-transfer group (ETM) having
XX different redox potentials. The invention is used for electronic
XX detection of nucleic acids, especially of substitutions (mismatches)
XX and single-nucleotide polymorphisms, e.g. for genotyping,
XX monitoring gene expression.
XX
SQ Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 16.1%; Score 108.8; DB 22; Length 936;
Best Local Similarity 0.7%; Pred. No. 5.1e-20;
Matches 5; Conservative 418; Mismatches 245; Indels 0; Gaps 0;

QY 9 ttccaatatcacccacccatcatgatgtcaagcacgcgcacacagtcggtttctg 68
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
59 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 gttattccctccagaatcactgcttcagaatatgagtcacagaacttcagccagacc 128
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
119 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129 ttccaactcaagcccttgcaaaatattgtctagaanaaagtgaatcttaggact 188
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
179 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
189 atccagatccgttttggaattatgacctttcttttgagatctctctttacactg 248
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
239 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
249 ttaaacacatcacagggttcccttattcttcagatcatalcattcggggctct 308
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
299 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
309 gtttggtcattaatctggagccttctaattgcagtgaaagaaaaaacacagaact 368
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
359 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
369 cgtataatatgagccgaataatgaaatctcttagtgccctgaggaataatcgtgaac 428
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
419 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
429 attctcctacattgtgttcactcctagatcaaaactacattgtgttatctcaccaa 488

```

```

DB 479 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
489 aatagtcagtgaaagcgtgtactgtctgtcttggaatttgatacattgact 548
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
539 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
549 ttcagcatatgaattatcattctctgccttctcctaatttgggtgccaccagag 608
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
599 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
609 gatgtgattgtgaacaatgtgttgactagcactgtgaataaatgtgttaata 668
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
659 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 669 aaaaaaaa 676
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
719 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
719 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
AAF58252/C
ID AAF58252 standard; DNA; 936 BP.
XX
XX AAF58252;
AC
XX 24-APR-2001 (first entry)
DT
XX
XX Oligonucleotide D1835.
DE
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
XX Synthetic.
XX
XX WO200107665-A2.
XX
XX 01-FEB-2001.
XX
XX 26-JUL-2000; 2000MO-US20476.
XX
XX 26-JUL-1999; 99US-0145695.
XX 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
XX hybridization assays, e.g. for genotyping, allowing repeat analyses on
XX a single surface.
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
XX acids each containing an electron-transfer group (ETM) having
XX different redox potentials. The invention is used for electronic
XX detection of nucleic acids, especially of substitutions (mismatches)
XX and single-nucleotide polymorphisms, e.g. for genotyping,
XX monitoring gene expression.
XX
SQ Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 16.1%; Score 108.8; DB 22; Length 936;
Best Local Similarity 0.7%; Pred. No. 5.1e-20;
Matches 5; Conservative 418; Mismatches 245; Indels 0; Gaps 0;

QY 9 ttccaatatcacccacccatcatgatgtcaagcacgcgcacacagtcggtttctg 68
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
729 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Oy 69 gtaattcctccagaatacactgcttcagaataatgagtcacagaacttcagccagacc 128
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 669 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 610
Oy 129 ttcttaactcaagcccttgcaaaaattattgctagaaaaatgaaactctaggact 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 609 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 550
Oy 189 atccagatcctgttggaataatgacacttctcttggaagtatctccttccacttg 248
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 549 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 490
Oy 249 ttaaaacatatcaaggttcccttataattcttcaagaatccacttgagctctc 308
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 430
Oy 309 gtttgatcatatctgagccttctccttaattgagtaagaagaacacacagaact 368
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 429 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 370
Oy 369 ctgataatattgagccgaataatgaaatctctgagtcgctgagacaatagctgaa 428
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 310
Oy 429 attcctccacacttggttccatctcctagatcaaaactcaatttggttattctc 488
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 309 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 250
Oy 489 aatagctagtgtaagcgtctgactgctgcttgctggaattgattgacttgact 548
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 249 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 190
Oy 549 ttcagcatattgaaatatactctcctgcttccacttgagggtgcacactagag 608
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 130
Oy 609 gatttgatctggaacaatgtgttgactagcacgtagaataaagtgtgtaaaata 668
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 70
Oy 669 aaaaaaa 676
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 wwwwwwwww 62
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 11

AF58254 AAF58254 standard; DNA; 936 BP.

AC AAF58254;

XX 24-APR-2001 (first entry)

DE Oligonucleotide D1875.

XX Electon-transfer group; ETM; mismatch; genotyping;

KM gene expression; ss.

OS Synthetic.

PN WO200107665-A2.

PD 01-FEB-2001.

PF 26-JUL-2000; 2000WO-US20476.

PR 26-JUL-1999; 99US-0145695.

PA 17-MAR-2000; 2000US-0190259.

PI (CLIN-) CLINICAL MICRO SENSORS INC.

```

XX WPI: 2001-159728/16.
DR Nucleic acids containing electron-transfer group, useful as labels in
XX hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
PT Example 6; Page 127; 159pp; English.
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
XX monitoring gene expression.
SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match      16.1%; Score 108.8; DB 22; Length 936;
Best Local Similarity 0.7%; Pred. No. 5.1e-20;
Matches 5; Conservative 418; Mismatches 245; Indels 0; Gaps 0;

Oy 9 ttcaaatatcacccgacacatcatgatacagacccgacacagtcggttctc 68
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 118
Oy 69 gtaattcctccagaatacactgcttcagaataatgagtcacagaacttcagccagacc 128
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 178
Oy 129 ttccaactcaagcccttcgcaaaaattattgctagaaaaatgaaactctaggact 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 238
Oy 189 atccaagatccgttggaataatgacacttctcttggaagtatcctccttcaacttg 248
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 229 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 298
Oy 249 ttaaaacatatcaaggttcccttataattcttcaaggaataccactcgggact 308
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 358
Oy 309 gtttgatcatatctgagccttccactaattgcagtgaaagaacacacagaact 368
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 418
Oy 369 ctgataataggccgaataatgaacttctcttgctgcccagagaagaatagctgaatc 428
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 478
Oy 429 attcctccacattggttccatctagatcaaaactcaatttggttattctcacc 488
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 538
Oy 489 aatagctagtgtaagcgtctgactgctccttggaatttgattgacttgact 548
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 539 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 598
Oy 549 ttcagcatattgaaatatactctcctgcttccacttgagggtgcacactagag 608
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 599 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 658
Oy 609 gatttgatctgacaatgtgttgactagcacgtgagataaagaatgtgtaaaata 668
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 659 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 718
Oy 669 aaaaaaa 676
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 719 wwwwwwwww 726
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 12

ID	AAF58254/C	standard; DNA; 936 BP.
XX	AAAF58254	
AC	AAF58254;	
XX		
DT	24-APR-2001	(first entry)
XX		
DE	Oligonucleotide D1875.	
XX		
KW	Electron-transfer group; ETW; mismatch; genotyping;	
KM	gene expression; ss.	
XX		
OS	Synthetic.	
XX		
PN	W0200107665-A2.	
XX		
PD	01-FEB-2001.	
XX		
PF	26-JUL-2000; 2000MO-US20476.	
XX		
XX	26-JUL-1999; 99US-0145695.	
XX	17-MAR-2000; 2000US-0190259.	
PA	(CLIN-) CLINICAL MICRO SENSORS INC.	
XX		
PI	umex RM;	
XX		
DR	WPI; 2001-159728/16.	
XX		
PT	Nucleic acids containing electron-transfer group, useful as labels in	
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on	
PT	a single surface	
XX		
PS	Example 6; Page 127; 159pp; English.	
XX		
CC	The present invention relates to a composition comprising two nucleic	
CC	acids each containing an electron-transfer group (ETW) having	
CC	different redox potentials. The invention is used for electronic	
CC	detection of nucleic acids, especially of substitutions (mismatches)	
CC	and single-nucleotide polymorphisms, e.g. for genotyping,	
CC	monitoring gene expression.	
XX		
SQ	Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;	
DB		
729	tttaaatcttcaccgcagaccatcatgatgattcaagcagccagacagtcgcgtgttctcg	68
DB	9	
YY	69 gttattccctcagaatacactgcttcagaataatgatlccacagaacttcagccagacc	128
DB	669	
OY	129 ttccaactcaagccctctgcgaataatttgcctcagaataatgaataatcttaggact	188
DB	609	
OY	189 atccagatccgtttggaataatgaactttctcttggaatatacttcctttcaacttg	248
DB	549	
OY	249 ttaaacatatccaaagttcccttataattcttcagatatccatcttgaggctct	308
DB	489	
OY	309 gttttgtcatatcttgagccttcctaattgcagtgaaagaaacacagaaact	368
DB	429	

Dy 369 ctgtaaatattgagccgaataatgaactctttagtgacctggagaaagtgtgaac 428

Dd 369 ww 310

Oy 429 attcctccacatttgtttcatcttcgatcaaaaacatttgtgttatctccaaca 488
::: : : ::|| ::: : : : : : : : : : : : : : : : :

Dd 309 www 250
ww

Oy 489 aatagtcagtgtaaagctgttaactgcgtgttcttggaaatttgatatgaat 548
::: :

Dd 249 www 190
www

Oy 549 ttosgcattatgaataatcattctctgccttctcaatltygggytgccactagaag 608
: :

Dd 189 www 130
www

Oy 609 gattgtgatttgacaacatgtgttgactcgcactgagagaataagatgtgttaata 668
: :

Dd 129 www 70
www

Oy 669 aaaaaaaa 676

Dd 69 :::::::::: 62
wwwwwwwww 62

Query Match	Best Local Similarity	Score	DB	Length
16.1%;	0.7%;	108.8;	DB 22;	936;
<p> AAF58257 AAF58257 standard; DNA; 936 BP. AAF58257; 24-APR-2001 (first entry) Oligonucleotide D1954. Electon-transfer group; ETM; mismatch; genotyping; gene expression; ss. Synthetic. WO200107665-A2. 01-FEB-2001. 26-JUL-2000; 2000WO-US20476. 26-JUL-1999; 99US-0145695. 17-MAR-2000; 2000US-0190259. (CLIN-) CLINICAL MICRO SENSORS INC. Umek RM; WPI; 2001-159728/16. Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface - Example 6; Page 127; 159pp; English. The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression. Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other; </p>				

[illegible]

PR	17-MAR-2000; 2000US-0190259.
XX	(CLIN-) CLINICAL MICHO SENSORS INC.
PA	
XX	Umek RM;
PI	
XX	WPI; 2001-159728/16.
DR	
XX	Nucleic acids containing electron-transfer group, useful as labels in
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT	a single surface
xx	
PS	Example 6; Page 127; 159pp; English.
CC	The present invention relates to a composition comprising two nucleic
CC	acids each containing an electron-transfer group (ETM) having
CC	different redox potentials. The invention is used for electronic
CC	detection of nucleic acids, especially of substitutions (mismatches)
CC	and single-nucleotide polymorphisms, e.g. for genotyping,
CC	monitoring gene expression.
xx	
SQ	Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;
Query Match	16.1%; Score 108.8; DB 22; Length 936;
Best Local Similarity	0.7%; Pred. No. 5,1e-20;
Matches	5; Conservative 418; Mismatches 245; Indels 0; Gaps 0;
OY	9 ttccaattatcacgcagccactgatgttcagaagcagcgacacagccagtgtttctg 68
DB	729 wmm 670
OY	69 gtaattccocagaatacactgtctcagaatalgagccacagaacttccagccagacc 128
DB	669 wmm 610
OY	129 tttcaacctcaaagccccttgcaaaatatlttgttagaaaaatgaaaacttaggact 188
DB	609 wmm 550
OY	189 alcagatccgtttgtgaatalgaccttctcttggagtatatcccttttcacttg 248
DB	549 wmm 490
OY	249 ttaaacaacatcacgaagttcccttattatcttctaagatcalcttgggctct 308
DB	489 wmm 430
OY	309 gtttgttcattaatctcggagccttcctaattgcatggaagaagaacacagaaact 368
DB	429 wmm 370
OY	369 ctgataatatlgagccgaalaatgaacttctctgtgcctcgaagacaatagctygaac 428
DB	369 wmm 310
OY	429 attccccacacattgtgttcatcctatagatacaaacatcattgtgttatctccaa 488
DB	309 wmm 250
OY	489 aatagtcagtglaagcgcttactgtcctgttctctgggaatttgattacattagact 548
DB	249 wmm 190
OY	549 ttacgcatlcatgaatatlcatcttctcgtcttcaatltygggtgcactcagag 608
DB	189 wmm 130
OY	609 gatctgatttgacaacatgttgtactagcacgtgagaaagaatggtttaaata 668
DB	129 wmm 70
OY	669 aaaaaaaaa 676

Db 69 MWMWMWMW 62

RESULT 15

ID AAF58259 standard; DNA; 936 BP.

XX AAF58259;

XX 24-APR-2001 (first entry)

DE Oligonucleotide D2004.

XX Electron-transfer group; ETM; mismatch; genotyping;

XX gene expression; ss.

XX Synthetic.

XX WO200107665-A2.

XX 01-FEB-2001.

XX 26-JUL-2000; 2000WO-US20476.

XX 26-JUL-1999; 99US-0145695.

XX 17-MAR-2000; 2000US-0190259.

XX (CLIN-) CLINICAL MICRO SENSORS INC.

XX Umeek RM;

XX WPI; 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in

XX hybridization assays, e.g. for genotyping, allowing repeat analyses on

XX a single surface

XX Example 6; Page 128; 159pp; English.

XX The present invention relates to a composition comprising two nucleic

XX acids each containing an electron-transfer group (ETM) having

XX different redox potentials. The invention is used for electronic

XX detection of nucleic acids, especially of substitutions (mismatches)

XX and single-nucleotide polymorphisms, e.g. for genotyping,

XX monitoring gene expression.

XX Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

XX Query Match 16.1%; Score 108.8; DB 22; Length 936;

XX Best Local Similarity 0.7%; Pred. No. 5.1e-20; Mismatches 245; Indels 0; Gaps 0;

XX Matches 5; Conservative 418; Mismatches 245; Indels 0; Gaps 0;

XX 9 ttcaaatatcacgcacacatcatgatgattcaagcaccgacacagtcggtttctg 68

XX 59 www. 118

XX 69 gtattccctcagaactcgtcgaataatgagtcacagaaacttccagccagacc 128

XX 119 www. 178

XX 129 tttaactcaagcccttgcaaaatattgtctagaataatgaaatcttaggact 188

XX 179 www. 238

XX 189 atccagatcctgttggaattgacattcttcttgagttacttccctttcacctg 248

XX 239 www. 298

XX 249 ttaaacatataccaagttcccttataattcttccagatataccatctcgggctct 308

XX 299 www. 358

QY 309 gtttgttcaattatctcgtgagccttccttaattgacgtgaaagaaaccacagaact 368

Db 359 www. 418

QY 369 ctgataatattgacgcgaataatgaatcttcttagtgccctgagagaatagctggaatc 428

Db 419 www. 478

QY 429 attctccacatttggttcaatccctagatcaaaactacatttggttattctcaccaa 488

Db 479 www. 538

QY 489 aataagtcagtgaagctgttactgttccgttcttgggaatttgatataatgagact 548

Db 539 www. 598

QY 549 ttcaagcattatgaattatcatctctcgtccttcccaatttgggggtgccaactcagag 608

Db 599 www. 658

QY 609 gattgtatttgaaacattgttgaactagacagtggagaataagatgtgttaataa 668

Db 659 www. 718

QY 669 aaaaaaa 676

Db 719 www. 726

Search completed: November 29, 2001, 00:23:58
Job time: 4315 sec

Fri Nov 30 11:02:22 2001

us-09-735-712-9.rng

Page 12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 00:20:23 ; Search time 83.78 Seconds
(without alignments)
1827.395 Million cell updates/sec

Title: US-09-735-712-9
Perfect score: 676
Sequence: 1 tcatctccttcaattatc.....tgyttaaataaaaaaa 676

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.4	12.8	1669	2	US-08-916-902A-2
2	86.4	12.8	1669	2	US-09-213-389-2
3	63.2	9.3	1661	1	US-08-318-492-3
4	63.2	9.3	1661	1	US-08-707-340-3
5	63.2	9.3	1661	2	US-08-994-578-3
6	49.2	7.3	2545	1	US-07-869-933-22
7	49.2	7.3	2545	4	US-09-103-663-22
8	42.6	6.3	7218	1	US-08-232-463-14
9	34.6	5.1	246240	2	US-08-724-394A-20
10	34.6	5.1	246240	2	US-08-724-394A-21
11	34.6	5.1	246240	2	US-08-724-394A-22
12	33.6	5.0	1711	1	US-08-568-147B-1
13	33.4	4.9	6669	3	US-09-212-971-5
14	33.4	4.9	6669	3	US-08-800-929A-5
15	33.4	4.9	6669	3	US-09-617-053A-5
16	32.2	4.8	3947	4	US-08-975-762-47
17	32.2	4.8	3947	4	US-09-295-028-47
18	32.2	4.8	3947	4	US-09-106-582-47
19	32	4.7	8920	2	US-08-446-855A-1
20	32	4.7	8920	2	US-09-150-741-1
21	32	4.7	9793	1	US-08-470-202-56
22	32	4.7	9793	1	US-08-471-770-56
23	32	4.7	9793	2	US-08-468-059-56
24	32	4.7	9793	4	US-09-109-916-56
25	31.8	4.7	1400	4	US-09-041-718-1
26	31.6	4.7	4171	1	US-08-308-881-5
27	31.6	4.7	4171	1	US-09-058-263-5

c 28	31.6	4.7	4171	2	US-09-059-099-5	Sequence 5, Appl1
c 29	31.6	4.7	4171	5	US-09-058-264-5	Sequence 5, Appl1
c 30	31.6	4.7	4171	5	PCT-US95-06530-5	Sequence 5, Appl1
c 31	31.4	4.6	3393	4	US-09-104-324B-1	Sequence 1, Appl1
c 32	31.4	4.6	3393	4	US-09-162-713-1	Sequence 1, Appl1
c 33	31.4	4.6	3560	4	US-09-081-345-1	Sequence 1, Appl1
c 34	31.4	4.6	4291	2	US-08-417-210A-81	Sequence 81, Appl1
c 35	31.4	4.6	4517	4	US-09-140-804-9	Sequence 9, Appl1
c 36	31	4.6	785	4	US-09-276-141-11	Sequence 11, Appl1
c 37	31	4.6	1902	4	US-08-793-044-4	Sequence 4, Appl1
c 38	31	4.6	2839	4	US-09-061-702-1	Sequence 1, Appl1
c 39	31	4.6	2945	3	US-09-058-489-85	Sequence 85, Appl1
c 40	31	4.6	9439	3	US-09-058-489-89	Sequence 89, Appl1
c 41	30.8	4.6	669	5	PCT-US95-13335-2	Sequence 2, Appl1
c 42	30.8	4.6	2662	3	US-08-750-357-8	Sequence 8, Appl1
c 43	30.8	4.6	3198	4	US-08-842-306B-48	Sequence 48, Appl1
c 44	30.8	4.6	3198	4	US-08-838-973B-48	Sequence 48, Appl1
c 45	30.8	4.6	3466	1	US-08-468-036-38	Sequence 38, Appl1

ALIGNMENTS

RESULT 1
US-08-916-902A-2
Sequence 2, Application US/08916902A
Patent No. 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,902A
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 927955
US-08-916-902A-2


```

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,492
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..741
US-08-318-492-3

```

```

Query Match          9.3%; Score 63.2; DB 1; Length 1661;
Best Local Similarity 50.9%; Pred. No. 1.7e-08;
Matches 239; Conservative 0; Mismatches 213; Indels 18; Gaps 3;

```

```

QY 167 aaaaatgaatcttaaggagataccagatccgtgttggaaatgaacttcttttg 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 234 AAAATTACAAAGTTCTGGGGCCATCCAGATCCGAATGCAGCAATGATCTGGCTTGGG 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 227 agt-----tacttcctttcaacctgttaaacacataccaagttccctttat 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 294 TGTCTTCTGGGTCTCTGCAATACCCATACACACTTCCAAAAGCACTTCTTCTTCAC 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 278 attcttcagagatataccatctctgggctctgtttgttcataatctcgagccctct 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 CTTCACAGAGCTACCCGATTTGGGGTCTGTGTTTCTGTAGTTCAGGAACCTTGTC 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 aattgcagtgaagaagaaacacagaaactctgataataattgagcgaataatgaatc 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 414 TGTGTAGCAGAGGATAAACCCCAAGAACATGATACAGAACAGTTTGGATGAAAT 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 398 tcttagtgccctgagagcaatagctggaatcatctccctcaattggtttcatcctaga 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 474 TGCACGTCTACAAATGGCACTAGTGGGGAGCTGCTTTCTCTCAATAATATGCACTTAA 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
458 tcaaaactacatttgygtat-tctcaacaaatagtcagtgtaagctgttact---- 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 534 TATCCAGTCATTAAGAGAGTTGTCACTCTTCATCAGAGTACCGGACCTATGCAATTAAT 593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 513 ---gtcctgtctctgggaatttgattacattgacttgcagcatattgaattat 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 594 GGGCTCCATATCAAAATGGCATGTGCTCTACTGCTGATTCACCTTGCTGGAATTATG 653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 569 cttctctgccttctcaatttgggtgctcaccaccagagaggttggatt 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 654 GGTAACTACTCTACCATATGCCATGTGTGCAATGCCAAATGCTGTGAATT 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 4
US-08-707-340-3
Sequence 3, Application US/08/707340
Patent No. 5705615
GENERAL INFORMATION:
APPLICANT: Iim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Leilas, Jean-Michel
TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
TITLE OF INVENTION: ASSAYS
NUMBER OF SEQUENCES: 4

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/707,340
FILING DATE: 03-SEP-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 97..738
US-08-707-340-3

```

```

Query Match          9.3%; Score 63.2; DB 1; Length 1661;
Best Local Similarity 50.9%; Pred. No. 1.7e-08;
Matches 239; Conservative 0; Mismatches 213; Indels 18; Gaps 3;

```

```

QY 167 aaaaatgaatcttaaggagataccagatccgtgttggaaatgaacttcttttg 226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 234 AAAATTACAAAGTTCTGGGGCCATCCAGATCCGAATGCAGCAATGATCTGGCTTGGG 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 227 agt-----tacttcctttcaacctgttaaacacataccaagttccctttat 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 294 TGTCTTCTGGGTCTCTGCAATACCCATACACACTTCCAAAAGCACTTCTTCTTCAC 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 278 attcttcagagatataccatctctgggctctgtttgttcataatctcgagccctct 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 CTTCACAGAGCTACCCGATTTGGGGTCTGTGTTTCTGTAGTTCAGGAACCTTGTC 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 aattgcagtgaagaagaaacacagaaactctgataataattgagcgaataatgaatc 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 414 TGTGTAGCAGAGGATAAACCCCAAGAACATGATACAGAACAGTTTGGATGAAAT 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 398 tcttagtgccctgagagcaatagctggaatcatctccctcaattggtttcatcctaga 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 474 TGCACGTCTACAAATGGCACTAGTGGGGAGCTGCTTTCTCTCAATAATATGCACTTAA 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 458 tcaaaactacatttgygtat-tctcaacaaatagtcagtgtaagctgttact---- 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 534 TATCCAGTCATTAAGAGAGTTGTCACTCTTCATCAGAGTACCGGACCTATGCAATTAAT 593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 513 ---gtcctgtctctgggaatttgattacattgacttgcagcatattgaattat 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 594 GGGCTCCATATCAAAATGGCATGTGCTCTACTGCTGATTCACCTTGCTGGAATTATG 653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 569 cattctctgccttctcaatttgggtgccactcagagattgtgatt 618
 | | | | | | | | | | | | | | | | | | | | | |
 Db 654 CGTAACATCTCTACCATAGCCATGTGGCAATGCAAACTGCTGTAAT 703

RESULT 5
US-08-994-

Sequence 3, Application US/08994578
Patent No. 5972688
GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Leilas, Jean-Michel
TITLE OF INVENTION: HTMA METHODS OF TREATMENT AND ASSAYS
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,578
FILING DATE: December 19, 1997
CLASSIFICATION: 433
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/707,340
FILING DATE: 03-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: Bih94-03A22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 3:
LENGTH: 1661 base pairs
type: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 97..738
US-08-994-578-3

Query Match	9.3%	Score 63.2	DB 2	Length 1661
Best Local Similarity	50.9%	Pred. No. 1.7e-08		
Matches 239	Conservative 0	Mismatches 213	Indels 18	Gaps 3

Qy 227 agt-----tatcttcctttcaccttgtaaacatatcaaglttccttat 277
||| | ||| | | | | | | |
Db 294 TGTCTTTCTGGGTCCTTGCATACCATACCACTCCAAAAGCATTCTTTCTTCAC 353

QY	278	atctcttccagagatccatctctgggctcggctttctcaatctatcttgagcccttc	337
Db	354	CTTCTACACAGGCTACCCGGATTTGGGGCGTGCTTTTCTGTAGTTCAAGAACCTTGTC	413
QY	338	aattgcggtgaagaagaaacccaagaactctgtatataattgagccggaataatgaatc	397
Db	414	TGTTGTACACAGGATTAACCCCAAGAACACTGGATTCAGAACACTTTTGGAAATGAACAT	473
QY	398	tcttagtgccttgagagcaatagcttggaatcatctctccatcttgatcttggttccatga	457
Db	474	TGCCAGTGCTCAATTTGCATAGTGGGGACGCTTTTCTGCATTAATATTAGCAGTTAA	533
QY	458	tcaaatcaactttgtgttat-ctccaccaaaatagtcggtgaagcgttact----	512
Db	534	TATCCAGTCATTTAAGGAGTTGTACTCTTCTACAGACTACCCGAGCTTAATGACATTTACAT	593
QY	513	---gtcctgtctcttgggaatttgaatcaattgaattgaattgaacttcaagcatatgaattat	568
Db	594	GGGCTCATATCAATAGGATGGTGTCTACTACTGATTTCTCACTTGGTGGAATTATG	653
QY	569	cattctctgccttcccaatttgggttgcgaactcagagatctgtgat	618
Db	654	CGTAACTATCTTCTACCAATTAGCCATGTGTGTGAATGCAACAACTGCTGTAAAT	703

RESULT 6
HS-07-869-

Sequence 22, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
TITLE OF INVENTION: IMUNOGLOBULIN
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/154 NIND
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 2545 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 46..786
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 46..54
FEATURE:

NAME/KEY: mat.peptide
LOCATION: 55..786
US-07-869-933-22

Query Match 7.3%; Score 49.2; DB 1; Length 2545;
Best Local Similarity 52.3%; Pred. No. 0.00015;
Matches 158; Conservative 0; Mismatches 138; Indels 6; Gaps 2;

QY 133 caactaaagccctgcgaataatcttgcagaaataatgaaatccttaggactatcc 192
DB 188 CACACAGAGAGACATGCGATCTTTTGAAAGAAAGAGTGGAGTTCGCGCGTAACCC 247
QY 193 agatcgtttggaatgaaccccttcttttgagttatcttcccttccactt----- 247
DB 248 AAGTTCTGTTGGTTGATATGCTTTGTTTGAAACAGTTGCTCCTCACACTCCAGA 307
QY 248 gttaaacatatccaaagtttccctta-tattcttcagagataccatctcgggct 306
DB 308 CTTCAGACTTTGACGAGAGAGTCTTTTATATAGACAGAGCTACCATCTCGGGTG 367
QY 307 cgttttgcattaatcttgagcccttccataatgcagtgaaagaaacacagaa 366
DB 368 CAGTCTGTTGTTGTTGCTGATTTTGTCAATATGTCGAAAGAAACACACACTGT 427
QY 367 cctgataatgatgcgaataatgaatcttcttagtgccctgagagcaatagctgaa 426
DB 428 ATCTGTGAGAGGACCTCGGAGCAACATTTCTCAGACATCGCTGCGCTTG6GGG 487
QY 427 tc 428
DB 488 TC 489

RESULT 7

US-09-103-663-22
Sequence 22, Application US/09103663D
Patent No. 6171803
GENERAL INFORMATION:
APPLICANT: Kinet et al.
TITLE OF INVENTION: Isolation, characterization, and use of the human beta
TITLE OF INVENTION: subunit of the high affinity receptor for
FILE REFERENCE: 50490
CURRENT APPLICATION NUMBER: US/09/103,663D
EARLIER FILING DATE: 1998-06-23
EARLIER APPLICATION NUMBER: 07/869,933
NUMBER OF SEQ ID NOS: 35
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 22
LENGTH: 2545
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: sig.peptide
LOCATION: (46)..(54)
FEATURE:
NAME/KEY: CDS
LOCATION: (46)..(786)
US-09-103-663-22

Query Match 7.3%; Score 49.2; DB 4; Length 2545;
Best Local Similarity 52.3%; Pred. No. 0.00015;
Matches 158; Conservative 0; Mismatches 138; Indels 6; Gaps 2;

QY 133 caactaaagccctgcgaataatcttgcagaaataatgaaatccttaggactatcc 192
DB 188 caactaaagccctgcgaataatcttgcagaaataatgaaatccttaggactatcc 247
QY 193 agatcgtttggaatgaaccccttcttttgagttatcttcccttccactt----- 247

DB 248 aagttcgtttggtttgatagtccttcttttggaacagttgctgcacacaccaga 307
QY 248 gttaaacatatccaaagtttccctta-tattcttcagagataccatctcgggct 306
DB 308 cttagactttgagcagagagttcttttatatagacagctacacattcttggtgtg 367
QY 307 cgttttgcattaatcttgagcccttccataatgcagtgaaagaaacacagaa 366
DB 368 cagtcgtttggtttgctgagattttgtcaattatgtccgaaagaaacacactgt 427
QY 367 cctgataatgatgcgaataatgaatcttcttagtgccctgagagcaatagctgaa 426
DB 428 atctgttgagagcagccttgagagcaacattgtcaagcagcatcgtcagagttggga 487
QY 427 tc 428
DB 488 TC 489

RESULT 8

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F15
US-08-232-463-14

Query Match 6.3%; Score 42.6; DB 1; Length 7218;
Best Local Similarity 6.6%; Pred. No. 0.015;
Matches 27; Conservative 204; Mismatches 178; Indels 0; Gaps 0;

```

QY 1 tcatctccttcaaatatcaccgacacatcagatgaagcaccgacacagtcg 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1094 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1153
QY 61 tcttctgctatctcctcgaagaactcgtctcgaatagccacgaacttcag 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1154 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1213
QY 121 ccagcacttccaactcaagccctcgaacaaatattctcgaacaaatgaatc 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1214 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1273
QY 181 taggacatccagactcgtctggaatagaccttctcttgagatcctcctt 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1274 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1333
QY 241 tcaacttgaacacatcgaagtcctccttatacttctcgaatatactct 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1334 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1393
QY 301 ggggctctgttctgctatcctgaagcctcctcctcctcctcagtgaaagaaaccca 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1394 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1453
QY 361 cagaacctgataatattgagccgacgaataatcctctctctagtcct 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1454 CTTTAACTTGATGATAGTAAATTAACAGTGAAGCTGCTACATCCG 1502

```

RESULT 9

```

US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant

```

```

; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note="HLA-H.CONTIG"
US-08-724-394A-20

```

```

Query Match 5.1%; Score 34.6; DB 2; Length 246240;
Best Local Similarity 55.4%; Pred. No. 9.2;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```

```

QY 481 ctccacaaatagtcagtgtaagctgtactcgtctccttggaatttgatcat 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36690 CTCCTCAGATCTCTATTATTAATTTATTTGCTGTGTAAGTAACACAACTG 36749
QY 541 tgatcttcagcatcttgatattcattcctcgtccttcctcctcctcctcctc 600
   ||||| ||| ||| ||| ||||| ||||| ||| ||| ||| ||| ||| |||
Db 36750 TGATGCTTGAACACACAAATTAATTTATTTACAAATTCYAAAGCTCAGAGTCC 36809
QY 601 a 601
Db 36810 A 36810

```

RESULT 10

```

US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..246240

```



```

LOCATION: 1:121040
; OTHER INFORMATION: /note= "HLA-H.CONFIG"
US-08-724-394A-22

```

Query Match	Best Local Similar
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Query Match	5.0%;	Score 33.6;	DB 1;	length 1711;
Best Local Similarity	47.6%;	Pred. No. 2.7;		
Matches	99;	Conservative	0;	Mismatches 109;
			Indels	0;
			Gaps	0;


```

; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPs AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009003
; CURRENT APPLICATION NUMBER: US/09/617, 053A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/800,929
; PRIOR FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (3677)...(3951)
; OTHER INFORMATION: n can be any nucleotide
US-09-617-053A-5

```

```

Query Match      4.9%; Score 33.4; DB 4; Length 6669;
Best Local Similarity 53.4%; Pred. No. 5.2;
Matches 70; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 160 ttgctagaaatgaataatcttagggactatccagatccctgttggatattgaccttt 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2946 TGGCTGTAAATATGACCTTCAAGATTTTAAATTTTGTGCCCTCAAGAACTAAC 2887
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 ctttggagttatcttccttccacttgccttgaataacataccaaggttcccttata 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2886 ATTGTAGAGTTATTCACACTATCACCCCTCGAAGAACATTTTGACTTTGTTTCTAA 2827
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 280 ttcttcaaga 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2826 ATGTTCTGAA 2816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: November 29, 2001, 00:23:11
Job time: 9823 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 00:18:42 ; Search time 1970.99 Seconds

(without alignments)
3685.532 Million cell updates/sec

Title: US-09-735-712-9

Perfect score: 676

Sequence: 1 tcatcccttcaatrac.....tgcgttaataaaaaaa 676

Scoring table: IDENTITY_NUC

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: qb_est1:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inu:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	503.2	74.4	508	10	A1149899 gf43h06.x
C 2	482.6	71.4	516	10	AA436088 zu03a08.r
C 3	379	56.1	387	10	AA416972 zt94h05.s
C 4	375.6	55.6	382	10	AA435988 zu03a08.s
C 5	325.2	48.1	394	10	AA758635 ah67b04.s
C 6	292.4	43.3	484	10	BE107659 UI-R-BP1-
C 7	274.8	40.7	415	10	AA470059 zt94h05.r
C 8	274	40.5	482	10	A1002083 ct38h02.s
C 9	267.4	39.6	410	10	AA411806 zt67a03.s
C 10	245.4	36.3	389	10	AA781801 at58g01.s
C 11	148.6	22.0	538	13	AQ108532 CIT-HSP-2
C 12	148	21.9	281	10	AA707529 ah41a12.s

C 13	141	20.9	332	11	BF319786	BF319786 uy63h11.x
C 14	131.8	19.5	601	10	BE638317	BE638317 EST000003
C 15	126.2	18.7	224	10	BE638325	BE638325 EST000022
C 16	86.4	12.8	1071	10	AL544561	AL544561 AL544561
C 17	84	12.4	793	11	BG484817	BG484817 602505511
C 18	83	12.3	218	10	BE638321	BE638321 EST00013
C 19	82.2	12.2	704	11	BE671626	BE671626 602592934
C 20	80.8	12.0	517	11	BE664654	BE664654 DRAGC07
C 21	80.4	11.9	508	10	AA234138	AA234138 zt51b05.r
C 22	80.4	11.9	562	10	AA418443	AA418443 zv92e05.r
C 23	76.2	11.3	857	11	BG720182	BG720182 602692226
C 24	70.6	10.4	421	13	BE6842	BE6842 R0C111-26F1
C 25	63.6	9.4	689	11	BG719295	BG719295 602690369
C 26	63.4	9.4	899	11	BE6720623	BE6720623 602691889
C 27	61.2	9.1	887	10	AL531049	AL531049 AL531049
C 28	59.4	8.8	715	11	BE538851	BE538851 602568296
C 29	55.6	8.2	579	10	BE513276	BE513276 601315340
C 30	55.6	8.2	947	12	AK008652	AK008652 Mus muscu
C 31	55.4	8.2	470	10	AI807884	AI807884 wF43q11.x
C 32	55	8.1	528	10	BE675149	BE675149 7f03d02.x
C 33	55	8.1	904	11	BE056859	BE056859 7k10g10.x
C 34	54.6	8.1	697	11	BE611537	BE611537 602613441
C 35	54.6	8.1	766	11	BE611537	BE611537 602613441
C 36	54.4	8.0	1504	12	AK003110	AK003110 Mus muscu
C 37	53	7.8	1184	12	AK017928	AK017928 Mus muscu
C 38	52.6	7.8	515	11	BE434337	BE434337 602506331
C 39	52.6	7.8	594	10	AV717594	AV717594 AV717594
C 40	52.6	7.8	597	10	AV716310	AV716310 AV716310
C 41	52.6	7.8	619	10	AV661743	AV661743 AV661743
C 42	52.6	7.8	680	11	BE546745	BE546745 602574115
C 43	52.6	7.8	735	10	AV715678	AV715678 AV715678
C 44	52.4	7.8	504	11	BE522968	BE522968 UI-R-C2P-
C 45	52.4	7.8	755	10	AI950360	AI950360 wp10b06.x

ALIGNMENTS

RESULT 1
A1149899/c
LOCUS gf43h06.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1752827
DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;
mRNA sequence.

ACCESSION A1149899
VERSION A1149899.1 GI:3678368
KEYWORDS EST.

ORGANISM human
SOURCE human
UNPUBLISHED (1997)

REFERENCE
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaapb@remail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldó
Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILN at:
www.bio.linnl.gov/bdtp/Image/Image.html
Insert Length: 742 Std Error: 0.00
Seq primer: -40ml3 fwd. EF from Amersham
High quality sequence stop: 455.

FEATURES

Source
1..508
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1752827"

```

/clone.lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
Inc., and primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGGCCGCCCAATTTTATTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      186 a      91 c      89 g      142 t
ORIGIN

```

```

Query Match      74.4%; Score 503.2; DB 10; Length 508;
Best Local Similarity 99.4%; Pred. No. 4.3e-102;
Matches 505; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 165 agaaaaaagaaatccttaggagactacagatcctgttggaataagaccttctt 224
    |||||||
Db 508 AGAAAAAGAAATCTTAGGGACATCCATGTTGGAATATGACCTTTCTTTT 449

QY 225 ggaattatcttcttcttcttcttcttcttcttcttcttcttcttcttctt 284
    |||||||
Db 448 GGAGTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 389

QY 285 tcagatatcatctctgagagctctgttcttcttcttcttcttcttcttcttca 344
    |||||||
Db 388 TCAGATATTCATCTGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 329

QY 345 gtgaaagaaacacacagaaactctgataatctgacgaataatgaaatcttct 404
    |||||||
Db 328 GTGAAAAAGAAACACAGAAACCTGATTAATATGACCCGAATATGAAATTTCT 269

QY 405 gccctgagagaaatagctggaatcttcttcttcttcttcttcttcttcttca 464
    |||||||
Db 268 GCCCTGGAGCAATAGCTGGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 209

QY 465 tacatttggtatcttccacaaatagctagctgtaagctgttcttcttcttctt 524
    |||||||
Db 208 TACATTTGTGTTATTTCTACCAAAATAGTCAAGTCAAGCTCTCTCTCTCTCT 149

QY 525 ggaatttgatcttcttcttcttcttcttcttcttcttcttcttcttctt 584
    |||||||
Db 148 GGAATTTTGATTTACATTCATGACCTTTCAGCATTTATTTGAAATTTCTCT 89

QY 585 tcaatttggggtgacctcagaagatgtgatttgaaatgttcttcttcttctt 644
    |||||||
Db 88 TCAATTTTGGGGTGCCTCAAGAGATGTGATTTGTGAACAATGTTGATGACCTG 29

QY 645 tgaataaagaatgtttaaataaaaaa 672
    |||||||
Db 28 TGAGAAATTAAGATGTGTTAAATAAAAA 1

RESULT 2
AA436088 516 bp mRNA EST 09-NOV-1997
LOCUS 2403a08.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730742
DEFINITION 5'. similar to SW:CD20_HUMAN P11836 B-LYMPHOCTYE ANTIGEN CD20 ''
mRNA sequence.
ACCESSION AA436088
VERSION AA436088.1 GI:2141002
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 516)
AUTHORS Hallier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Joat,S.,

```

```

TITLE
JOURNAL
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 733 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 482.

FEATURES
Source
1..516
/organism="Homo sapiens"
/db_xref="GDB:5927437"
/db_xref="taxon:9606"
/clone="IMAGE:730742"
/clone.lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
Inc., and primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGGCCGCCCAATTTTATTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      149 a      117 c      80 g      170 t
ORIGIN

Query Match      71.4%; Score 482.6; DB 10; Length 516;
Best Local Similarity 99.2%; Pred. No. 1.6e-97;
Matches 485; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tcaatccttccaatattacacgacacatcagatcagaacgcagacagctcgg 60
    |||||||
Db 20 TCATCTCTCTTCAATTAATATACCGACACATCAGATTAAGACCCACACATCCGG 79

QY 61 tgttctgtatcttccacgaatcactgcttgcgaatgagtcacagaaacttctag 120
    |||||||
Db 80 TGTCTTCTGTATTTCTCTCAAGAAATCACTGCTTCAGAAATATGATGCCAGAACTTTAG 139

QY 121 ccagaccttccaactcaaaagcccttgcacaaatatttgcctagaacaaatgaatct 180
    |||||||
Db 140 CCAGACCTTTTCAACTCAAAAGCCCTTGCAAAATATTGCTGAGAAATVGAATAATCT 199

QY 181 taggacataccagatcctgttggaaatagacatttcttggagatcttctctt 240
    |||||||
Db 200 TAGGACTATCCAGATCTCTGTTTGGAAATTAAGACCTTTCTTTGGAAATTAATCTCTT 259

QY 241 tcacctgttaaacacatccaaggttcccttatattcttcttcaagatccatct 300
    |||||||
Db 260 TCACCTTTTAAACCAATTCAGAGTTCCCTTATATTTCTTTCAGAGATATTCATTCT 319

QY 301 ggggctctgttcttcttcaatctctgagccttccatctgagtgagaaacacaa 360
    |||||||
Db 320 GGGGCTCTGTTTGTCTATTAATCTGGAGCCTTCCTTAATGAGTAAAGAAACCA 379

QY 361 cagaacactgataatattgagccgaataatgaaatctcttaagtcctgagcaatag 420
    |||||||
Db 380 CAGAAACTGTATATATATGAGCCGAATTAAGAAATTTCTTAGGCCCTGGAGCAATAG 439

QY 421 ctggaatcctctcactcattgttctcactcagatcaaaatcttcttcttcttctt 480
    |||||||

```

```

Db      440 CTGGAAATCATTCCTCAGATTGTTTCATCCTAGATCAAAATACATTGTGTTATT 499
QY      481 CTCACCAAA 489
Db      500 CTCACAAA 508

RESULT  3
LOCUS   AA416972/c      387 bp      mRNA      EST      09-NOV-1997
DEFINITION   zt94n05.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730041
              3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;
              mRNA sequence.
ACCESSION   AA416972
VERSION     AA416972
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 387)
AUTHORS    Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gelsel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            J., Moore,B., Schellenberg,K., Stepien,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
            WashU-NCI human EST Project
            Unpublished (1997)
JOURNAL
TITLE      Contact: Wilson RK
COMMENT     Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: estewatson.wustl.edu
            This clone is available royalty-free through LNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 656 Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham.
FEATURES
source     1..387
            /organism="Homo sapiens"
            /db_xref="GDB:5926570"
            /db_xref="taxon:9606"
            /clone="IMAGE:730041"
            /clone_1ib="Soares_testis_NHT"
            /sex="male"
            /lab_host="DH10B"
            /note="Vector: pRT73D-Pac (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
            was prepared from mRNA obtained from Clontech Laboratories
            , Inc., and primed with a Not I - oligo(dT) primer [5',
            TGTTCACATCTGAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3'].
            Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pRT73 vector. Library
            went through one round of normalization to Cot5, and was
            constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 138 a      77 c      67 g      105 t
ORIGIN
Query Match      56.1%; Score 379; DB 10; Length 387;
Best Local Similarity 98.7%; Pred. No. 2e-74;
Matches 382; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

Db      267 TAGTGCCCTGGAGCAATAGCTGGAAATCATTCCTCAGATTGTTTCATCCTAGATCA 208
QY      461 aaactacattgtgttatcttcaccaaatagtcagtgtaagcctgttactctgtt 520
Db      207 AAACATCATTTTGGTGTATTCCTCACCMAAATAGCTAGTGAAGCTGTACTCTCTGT 148
QY      521 ctgggaatttgattacattgacttcagcattatgaatttcaattcttcgcc 580
Db      147 CTGGGAATTTTGATTCATGATGACTTTCACATTTATGAAATTCATTTCTCTGCC 88
QY      581 ttctcaatttgggtgcacacacagagatttgatttgcaacaatgtgttgaactagc 640
Db      87 TTTCTCAATTTTGGGGTGCACACACAGAGATTGTGATTTGACAAATGTTGTGACTAGC 28
QY      641 actgtgagataaagaatgtgttaaat 667
Db      27 ACTGTGAGATAAAGATGTGTTAAAT 1

RESULT  4
LOCUS   AA435988/c      382 bp      mRNA      EST      09-NOV-1997
DEFINITION   zu03a08.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730742
              3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;
              mRNA sequence.
ACCESSION   AA435988
VERSION     AA435988
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 382)
AUTHORS    Hillier,L., Allen,M., Bowles,L., Dubuque,T., Gelsel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            J., Moore,B., Schellenberg,K., Stepien,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
            WashU-NCI human EST Project
            Unpublished (1997)
JOURNAL
TITLE      Contact: Wilson RK
COMMENT     Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: estewatson.wustl.edu
            This clone is available royalty-free through LNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Possible reversed clone: similarity on wrong strand
            Insert Length: 733 Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 280.
FEATURES
source     1..382
            /organism="Homo sapiens"
            /db_xref="GDB:5927437"
            /db_xref="taxon:9606"
            /clone="IMAGE:730742"
            /clone_1ib="Soares_testis_NHT"
            /sex="male"
            /lab_host="DH10B"
            /note="Vector: pRT73D-Pac (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
            was prepared from mRNA obtained from Clontech Laboratories
            , Inc., and primed with a Not I - oligo(dT) primer [5',
            TGTTCACATCTGAGTGGAGCGCGCCCAATTTTCTTTTCTTTT 3'].
            Double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pRT73 vector. Library
            went through one round of normalization to Cot5, and was
            constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 135 a      73 c      67 g      107 t
ORIGIN

```

Query Match 55.6%; Score 375.6; DB 10; Length 382;
 Best Local Similarity 99.0%; Pred. No. 1.1e-73;
 Matches 378; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

291 tatccattcgtggagctgtgtttgttcaatattctgagcctctcttaattgcaatgaa 350
 |||||||
 302 tatccattcgtggagctgtgtttgttcaatattctgagcctctcttaattgcaatgaa 323
 |||||||
 351 agaaaacacacagaaactctgataatattgagccgaataatgatactctttagtcctc 410
 |||||||
 322 agaaaacacacagaaactctgataatattgagccgaataatgatactctttagtcctc 263
 |||||||
 411 agagaataatagctggaatactctctcacaattgttcaatccttagatcaaaactacatt 470
 |||||||
 262 ggagcaataatagctggaatactctctcacaattgttcaatccttagatcaaaactacatt 203
 |||||||
 471 tttgtttcttcacacaaatagctggaatactctctcacaattgttcaatccttagatcaaa 530
 |||||||
 202 tttgtttcttcacacaaatagctggaatactctctcacaattgttcaatccttagatcaaa 143
 |||||||
 531 ttgattacattgattacattcagacattatgaaatcattcctcctgcttcccaatt 590
 |||||||
 142 ttgattacattgattacattcagacattatgaaatcattcctcctgcttcccaatt 83
 |||||||
 591 ttgggtgtccacacagagattgtgattgtgacaagtgttgttgaactgagactgtgagaa 650
 |||||||
 82 ttgggtgtccacacagagattgtgattgtgacaagtgttgttgaactgagactgtgagaa 23
 |||||||
 651 taagaatgtttaaataaaaa 672
 |||||||
 22 TAAAGATGTGTTAAATATATAA 1
 |||||||

RESULT 5
 AA758635/c 394 bp mRNA EST 29-DEC-1998
 LOCUS ab67b04.s1 Soares testis_NHT Homo sapiens cDNA clone 1320655 3'
 DEFINITION similar to SW:CD20_HUMAN F11836 B-LYMPHOCTE ANTIGEN CD20'', mRNA
 sequence.
 ACCESSION AA758635.1 GI:2806498
 VERSION AA758635
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 394)
 NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-remail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
 www-bio.llnl.gov/bdrrp/image/image.html
 Insert Length: 661 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 390.
 Location/Qualifiers
 1..394
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="1320655"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"

FEATURES

1..394
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="1320655"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTCAATCTGGAAGTGGAGGCGGCCCAATTTTGTGTTTGTGTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

Query Match 48.1%; Score 325.2; DB 10; Length 394;
 Best Local Similarity 95.7%; Pred. No. 1.8e-62;
 Matches 377; Conservative 0; Mismatches 13; Indels 4; Gaps 4;

281 tcttccagatattcattctgtggtctgtttgttcaatattctgagccttcccaatt 340
 |||||||
 394 tcttccagatattcattctgtggtctgtttgttcaatattctgagccttcccaatt 335
 |||||||
 341 tgcagtgaagaagaacacacagaaactctgataattgagccgaataatgaaactctc 400
 |||||||
 334 tgcagtgaagaagaacacacagaaactctgataattgagccgaataatgaaactctc 275
 |||||||
 401 tagtgccttgagagaaatagctggaatcattcctcacaattg-gttcattcctagac 459
 |||||||
 274 tagtgccttgagagaaatagctggaatcattcctcacaattg-gttcattcctagac 215
 |||||||
 460 aaactacattgtgttattctcaccaaatagtcagtgtaaggctgttaactgtcctgt 519
 |||||||
 214 aaactacattgtgttattctcaccaaatagtcagtgtaaggctgttaactgtcctgt 155
 |||||||
 520 tcttgggaatttgaatcattgacattcagcat-tattgaattcaat-ttctct 577
 |||||||
 154 tcttgggaatttgaatcattgacattcagcat-tattgaattcaat-ttctct 95
 |||||||
 578 gccctt-ctcaatttgggggtgcacctaagagattgtgattgaaacattgttgac 636
 |||||||
 94 gccctt-ctcaatttgggggtgcacctaagagattgtgattgaaacattgttgac 35
 |||||||
 637 tagcactgtgagaataagaatgtgttaaaataa 670
 |||||||
 34 TAGCACTGTGAGATTAAGATATGTTAAATATAA 1
 |||||||

RESULT 6
 BE107659/c 484 bp mRNA EST 13-JUN-2000
 LOCUS U1-R-BT1-ame-d-04-0-01.s1 U1-R-BT1 Rattus norvegicus cDNA clone
 DEFINITION U1-R-BT1-ame-d-04-0-01 3', mRNA sequence.
 ACCESSION BE107659
 VERSION BE107659.1 GI:8499769
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 484)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized testis library cDNA Library Preparation: M.B. Soares Lab clone distribution: clones will be available through Research Genetics (www.resgen.com)

Seq primer: M13 Forward

POLYA=Yes.

FEATURES

source

Location/Qualifiers

1..484

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-BT1-ame-d-04-0-UI"

/clone_lib="UI-R-BT1"

/dev_stage="adult"

/lab_host="DHI0B (Life Technologies)"

/note="Vector: pUT3D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; The library UI-R-BT1 is a subtracted library derived from a mixture of the following tissues: hippocampus, thalamus, mid-brain, medulla, corpus striatum, cerebral cortex and testis. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)

TAG_LIB=UI-R-BT1

TAG_TISSUE=testis

TAG_SEQ=ACCCAG

BASE COUNT 165 a 98 c 96 g 125 t

ORIGIN

Query Match 43.3%; Score 292.4; DB 10; Length 484;
Best Local Similarity 76.9%; Pred. No. 3.5e-55;
Matches 372; Conservative 0; Mismatches 106; Indels 6; Gaps 1;

DB 193 agatccgttgggaattgacattcttcttgaggtatcttcttccattcactgttaa 252
|||||
DB 484 AGATCCGTTGGGAATGAACTTCATTTGGAGTTGTTCCCTTTCACCTGTAA 425
|||||
DB 253 aacctaccgaagttcccttatattcttcaggatattccatctgggctctgtt 312
|||||
DB 424 ACCCATACCAAGTTCCTTTATATTTCTCTGAGATATCCCTTCTGGGCTCTGCT 365
|||||
DB 313 tttcattcaattctgagccttcctcaattgagtgaaagaaacacacagaactctga 372
|||||
DB 364 TGTTCATTAACCTGAGCTTCTGATTCGTGAAAAAGAAAACTACAGCACTTGA 305
|||||
DB 373 taatttgagccgaataatcttcttagtgccctgagagcaatagctgaatcattc 432
|||||
DB 304 TAAAAATGAGCGAGCGATGATTTACTTAGTGCCCTGGAGCAGACGCTGGAATCATTC 245
|||||
DB 433 tccatcatttggttcctcctgaagaaactacatttggtatctccacaaata 492
|||||
DB 244 TCCCTATATGTCCTTCTTAGATGAGAAATCATCTGTGGCTATTTCTCAGATGGA 185
|||||
DB 493 gtcaatgtgaagcgttactgtccgtctctggaatttgatcatgatgaattcca 552
|||||
DB 184 TTCAGTGGGTCTATACACTATTTGTTGGGATTTTATGTTATGATGAGATCTTCA 125
|||||
DB 553 gcatatgaattatcattctctgccttctcaatttggggtgcaactcagaagatt 612
|||||
DB 124 GCGTGTGCAACTGTCATTTCCCTTTCTTCATTTTGGGGTGTACTCAGAGAAA 65
|||||
DB 613 ggtatgtgaagaaatgtgtgacagacgttgagaaataagaagtggttaaaataaa 672
|||||
DB 64 GTGGGGG-----ATGCTGTGATATGACATGTACAGTAAACAGCAAAATTGAAAAAA 11
|||||
DB 673 aaaa 676

DB 10 AAAA 7

RESULT 7

AA470059 415 bp mRNA EST 09-NOV-1997
LOCUS zt94h05.r1 Soares testis NHT Homo sapiens cDNA clone IMAGE:730041
DEFINITION 5' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ; ;
mRNA sequence.

ACCESSION AA470059

VERSION AA470059.1 GI:2197368

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 415)

AUTHORS Hillier,L., Allam,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin

JOURNAL White,Y., Wylie,T., Waterston,R. and Wilson,R.

COMMENT MashU-NCI human EST Project

Unpublished (1997)

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert length: 656 Std Error: 0.00

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 393.

FEATURES

source

Location/Qualifiers

1..415

/organism="Homo sapiens"

/db_xref="GDB:5926570"

/db_xref="taxon:9606"

/clone="IMAGE:730041"

/clone_lib="Soares_testis_NHT"

/sex="male"

/lab_host="DHI0B"

/note="Vector: pUT3D-Pac (Pharmacia) with a modified

polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was prepared from mRNA obtained from Clontech Laboratories

, Inc., and primed with a Not I - oligo(dT) primer [5'

TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pUT3D vector. Library

went through one round of normalization to Cot5, and was

constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT

129 a 96 c 67 g 123 t

ORIGIN

Query Match 40.7%; Score 274.8; DB 10; Length 415;
Best Local Similarity 78.7%; Pred. No. 2.9e-51;
Matches 389; Conservative 0; Mismatches 2; Indels 103; Gaps 1;

DB 1 tcatctcttcaattatccacgacacatcatgatgaagacacgacacagctcgg 60
|||||
DB 25 TCAATCTCTTCAAAATATACACGACCATCATGATTCAGCAACGACACAGTCCGG 84
|||||
DB 61 tttctctgtaattctccacgaatacactgtctcagaatatgagtcacagaacttcag 120
|||||
DB 85 TGTTCCTGTGTAATTCCTCCAGAAATCACTGCTTCGATATGAGTCCACAGACTTCAG 144
|||||
DB 121 ccagacatttcaactcaagaagcccttgcaaaaattatctgtctagaaataatgaatct 180
|||||
DB 145 CCACGACCTTTCAACTCAAGCCCTTGCAAAAATATATTGCTAGAAAATGAAAATCT 204
|||||

```

QY 181 taggacatccagatccgtgttgaaatagacccttcttcttgaggtatcttccctt 240
    |||
Db 205 TAGG----- 208
QY 241 tcacctgtttaaaccatataccaagttcccttatacttcttcagagatccattct 300
    |||
Db 209 -----GGATATCCATTCT 221
QY 301 ggggctctgtttgttcattatctcagagccttcctaatctgacgtgaaagaacaa 360
    |||
Db 222 GGGGCTCTGTTTGTTCATTATCTGGAGCCTTCCTTAATTCAGAGTGAAGAAAGAACCA 281
QY 361 cagaactctgataatattgagccgaataatgaatcttcttagtgccttgagagaatag 420
    |||
Db 282 CAGAACTCTGATTAATATGAGCCGAATTAAGAAATTTCTTAGTCCCTGGAGACAATAG 341
QY 421 ctggaatcatcttctcctacattgtgttctcctcagatcacaactacattgtgtatt 480
    |||
Db 342 CTGGAATCATCTCTCCTACATTGTGTTCACTCCTAGATCAAACTACATTTGTGTTATT 401
QY 481 ctacccaataagt 494
    |||
Db 402 CTCACCAAAATAGT 415

```

```

RESULT 8
LOCUS AI002083 482 bp mRNA EST 27-AUG-1998
DEFINITION 013802.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1619019
ACCESSION 3, mRNA sequence.
VERSION AI002083
KEYWORDS AI002083.1 GI:3202120
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 482)
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@email.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/dbp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40m13 fwd. RT from Amersham
High quality sequence stop: 393.
Location/Qualifiers
1.482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1619019"
/clone_11b="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73d-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I oligo(dT) primer (5'
TGTTCACCAATCTAGAGTGGAGGCGCCCAATTTTGTGTTT 3'}.
Double-stranded cDNA was ligated to Eco RI adaptors.
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```

FEATURES

```

source
1.482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1619019"
/clone_11b="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73d-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I oligo(dT) primer (5'
TGTTCACCAATCTAGAGTGGAGGCGCCCAATTTTGTGTTT 3'}.
Double-stranded cDNA was ligated to Eco RI adaptors.
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```

```

BASE COUNT 157 a 85 c 99 g 141 t
ORIGIN
Query Match 40.5%; Score 274; DB 10; Length 482;
Best Local Similarity 71.9%; Pred. No. 4.3e-51;
Matches 482; Conservative 0; Mismatches 0; Indels 188; Gaps 2;

```

```

QY 1 tcattcccttcaaatatccacagacacatcagatctcagacagcagacagtcagg 60
    |||
Db 482 TCATCTCTCTTCAATATATACACGACACATCATGATCAGACACGACACAGTCCG 423
QY 61 tttcttgatttctcctcagaaatcactgcttcaagaatagttccagaaacttcag 120
    |||
Db 422 TGTTCCTGATTTCTCCCAAGAAATCACTGCTCAGATATGATGACACAGAACTTCAG 363
QY 121 ccagacatttcaactcaagagcccttgaaatattatgtcagaaaaatgaaatct 180
    |||
Db 362 CCAGACCTTTTCAACTCAAGCCCTTGGCAAAATATTTGCTAGAAAAATGAATATCT 303
QY 181 taggacatccagatccgtgttgaaatagacccttcttcttgaggtatcttccctt 240
    |||
Db 302 TAGG----- 297
QY 241 tcacctgtttaaaccatataccaagttcccttatacttcttcagagatccattct 300
    |||
Db 298 -----GGATATCCATTCT 286
QY 301 ggggctctgtttgttcattatctcagagccttcctaatctgacgtgaaagaacaa 360
    |||
Db 285 GGGGCTCTGTTTGTTCATTATCTGGAGCCTTCCTTAATTCAGTGAAGAAAGAACCA 226
QY 361 cagaactctgataatattgagccgaataatgaatcttcttagtgccttgagagaatag 420
    |||
Db 225 CAGAACTCT----- 214
QY 421 ctggaatcatcttctcctacattgtgttctcctcagatcacaactacattgtgtatt 480
    |||
Db 215 -----GATCAAACTACATTTGTGTTATT 191
QY 481 ctacccaataatagtcagtgtaagcgttctactgctcctgcttcttggaatttgatcat 540
    |||
Db 190 CTCACCAAAATAGTCAATGTAAGGCTGTACTGTCTGCTTCTTGGAATTTGATTACAT 131
QY 541 tgaatccttcagacattatgaatattatccttctccttctcaatttgggtgccc 600
    |||
Db 130 TGATGACTTTCAGCATTAATGAATTAATTCATTCTGCTCTTCTCAATTTGGGGTGGCC 71
QY 601 actaagagatttgatgtgaaacatggtgttctactagacactgtgagataaagatgtg 660
    |||
Db 70 ACTCAGAGGATTTGATTTGTAACAAATGTTGTGACACACTGTGAGAAATTAAGATGTG 11
QY 661 ttaatacaaa 670
    |||
Db 10 TTAATAATAAA 1

```

```

RESULT 9
LOCUS AA411806 410 bp mRNA EST 12-AUG-1997
DEFINITION 2t67a03.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:727372
ACCESSION 3, similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
VERSION AA411806
KEYWORDS mRNA sequence.
SOURCE EST.
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 410)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Maitra, M., Martin, J., Moore, B.

```

TITLE
JOURNAL
COMMENT
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Meterston, R. and Wilson, R.
Washu-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estelw@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 608 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 401.
Location/Qualifiers

FEATURES

source

1. 410
/organism="Homo sapiens"
/db_xref="GDB:5924284"
/db_xref="taxon:9606"
/clone="IMAGE:727372"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCCGCCCAATTTTCTTTTGT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 147 a 68 c 119 t
ORIGIN

Query Match 39.6%; Score 267.4; DB 10; Length 410;
Best Local Similarity 96.1%; Pred. No. 1.3e-49;
Matches 274; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 113 atttcaagcagcagccttcaaccacaagcccttgcaaaaattattgtcagaataat 172
|||||
DB 410 ACTTTACGCGACGACCTTTTCAACCAAGCCCTTGCAAAATATTATTTGCTAGAAAAAT 351
QY 173 gaaatcttaaggactatccagatccctgttggaattatgaccttcttcttgaggtat 232
|||||
DB 350 GAAAAATCTTAGGAGCTATCCAGATCTCTGTTGGAATTATGACCTTTCTTTGGAGTTAT 291
QY 233 ctctctttcaccttgtaaaacacatccaaagttcccttataattcttcaaggata 292
|||||
DB 290 CTTCCCTTTACCTTGTAACCAATCCAAAGTTCCCTTTATATTTCTTTCAAGGATA 231
QY 293 tccattctggagctctgtttgttcatatattcttgagccctcccaatgcagtaaaag 352
|||||
DB 230 TCCATTCTGGGCGCTCTGTTTCTCATATATCTGAGGCTTCCTAATTCGACGTAAGAAG 171
QY 353 aaaaaccacagaactctgataataattgagccgaataatgaactc 397
|||||
DB 170 AAAAAACACAGAACTCTGGGAATTTGTATGTACATTTGATGACTTT 126

RESULT 10
AA781801/c 389 bp mRNA EST 31-DEC-1998
LOCUS AA781801 a158901.s1 Soares_testis_NHT Homo sapiens cDNA clone 1375056 3'
DEFINITION similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ; mRNA
ACCESSION AA781801
VERSION AA781801.1 GI:2841132
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
cDNA library preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution by: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbbrp/image/image.html
Insert Length: 983 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 311.
Location/Qualifiers

FEATURES

source

1. 389
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1375056"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCCGCCCAATTTTCTTTTGT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 140 a 66 c 68 g 115 t
ORIGIN

Query Match 36.3%; Score 245.4; DB 10; Length 389;
Best Local Similarity 95.8%; Pred. No. 1e-44;
Matches 252; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 135 actcaaaagcccttgcanaaattatttctgtagaaaatgaaacttaggactatccag 194
|||||
DB 389 ACTCAAGCCCTTGCAAAATATTATTTGCTAGAAAAATGAAATCTTAGGACTATCCAG 330
QY 195 atccgtttggaattatgaccttcttcttggaattatcttcttcccttgtaaaa 254
|||||
DB 329 ATCCGTGTTGGAAATATGACCTTTCTTTGGAGTATCTTTCCCTTTGACCTGTTAAAA 270
QY 255 caatcccaaggttcccttataattcttcaagatatccattctgggctctgtttg 314
|||||
DB 269 CCATATCCAGAGTTCCCTTATATTCTTTACAGATATCCATTTCTGGGCGCTCTGTTTG 210
QY 315 ttcaatlaattctgagccttctctaattgcagtgaaaagaaaaacacagaaaacttgata 374
|||||
DB 209 TTCATTAAATTCGAGGCTTCTTAATTGCAAGTGAAGAAACACAGAAACTCTGGGA 150
QY 375 atattgagccgataatgaactc 397
|||||
DB 149 ATTTGATTACATTTGATGACTTT 127

RESULT 11
AQ108532 538 bp DNA GSS 29-AUG-1998
LOCUS AQ108532 CIT-HSP-237967.TP CIT-HSP Homo sapiens genomic clone 237967, DNA
DEFINITION
ACCESSION AQ108532
SEQUENCE.

VERSION AQ108532.1 GI:3485222
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 538)
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
 Venter,J.C.
 TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
 Map Building
 JOURNAL Unpublished (1998)
 COMMENT Other_GSSs: CIT-HSP-237967.TR
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: madams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tbdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.
 FEATURES
 source Location/Qualifiers
 1..538
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="237967"
 /clone_lib="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBelOBAC11; site_1: HindIII; site_2:
 HindIII"
 BASE COUNT 140 a 108 c 115 g 175 t
 ORIGIN
 Query Match 22.0%; Score 148.6; DB 13; Length 538;
 Best Local Similarity 94.5%; Pred. No. 3.3e-23;
 Matches 154; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 Oy 506 tcttactgcctgtcttctgggaatttgatcattcattgacattcattgaatt 565
 |||||
 Db 152 tcttcttcttcttcttctggaatttgatcattcattgacattcattgaatt 211
 |||||
 Oy 566 attcatttctgccttctcatttctgggtgcacacagagatttgattgagaac 625
 |||||
 Db 212 attcatttctgccttctcatttctgggtgcacacagagatttgattgagaac 271
 |||||
 Oy 626 atgtctgtactagacactgtgagaataaagatgtgttaataa 668
 |||||
 Db 272 atgtgttgcactagacactgtgagaataaagatgtgttaataa 314
 |||||
 RESULT 12
 AA707529 281 bp mRNA EST 12-JAN-1999
 LOCUS ab1a12.sl Soares_testis_NHT Homo sapiens cDNA clone 1292062 3',
 mRNA sequence.
 ACCESSION AA707529
 VERSION AA707529.1 GI:2717447
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 281)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs@mail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
 , Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LML at:
 www-bio.llnl.gov/bbrp/image/image.html
 Insert length: 456 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 218.
 FEATURES
 source Location/Qualifiers
 1..281
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="1292062"
 /clone_lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization to Cots, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 94 a 54 c 45 g 88 t
 ORIGIN
 Query Match 21.9%; Score 148; DB 10; Length 281;
 Best Local Similarity 100.0%; Pred. No. 5.2e-23;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 523 tgggaatttgattacattgacttgacttcagcattatgaattatcattctctgcct 582
 |||||
 Db 148 TGGGATTTTGTATTACATTGATGACTTTCACCATTTGATTTCTCTGCCTT 89
 |||||
 Oy 583 tctcaatttgggtgcacacagagatttgattgtgaacaattgttgaactagac 642
 |||||
 Db 88 TCTCATTTTGGGGGCCACACAGAGATTTGATTTGGAACAATGTTGACTAGCAC 29
 |||||
 Oy 643 ttgtgagaataagatgtgttaataa 670
 |||||
 Db 28 TGTGAGATTAAGAATGTTTAAATAA 1
 |||||
 RESULT 13
 BF319786 332 bp mRNA EST 29-DEC-2000
 LOCUS uy63h11.x1 McCarrey Eddy round spermatid Mus musculus cDNA clone
 IMAGE:3664293 3', mRNA sequence.
 ACCESSION BF319786
 VERSION BF319786.1 GI:11268601
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 332)
 AUTHORS Underwood,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person
 ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler
 ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.
 The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Mairra M/WashU-NCI Mouse EST Project 1999

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2001, 21:32:37 ; Search time 1789.67 Seconds

(without alignments)
4148.100 Million cell updates/sec

Title: US-09-735-712-7

Perfect score: 450

Sequence: 1 atgtatcaagcagcgcaca.....attgtgacaactgtgtga 450

Scoring table: IDENTITY_NUC

Searched: 1472140 seqs, 8248589755 residues

al number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_hg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_om:*

20: em_or:*

21: em_ov:*

22: em_pat:*

23: em_ph:*

24: em_pl:*

25: em_ro:*

26: em_sts:*

27: em_sy:*

28: em_un:*

29: em_vl:*

30: em_hgo_hum:*

31: em_hgo_inv:*

32: em_hgo_rnd:*

33: em_hgt_hum:*

34: em_hgt_inv:*

35: em_hgt_rnd:*

36: em_hgt_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	450	100.0	450	6	AX179787	AX179787 Sequence
2	347.4	77.2	603	6	AX179781	AX179781 Sequence
3	347.4	77.2	676	6	AX179789	AX179789 Sequence
4	347.4	77.2	691	9	AB013103	AB013103 Homo sapi
5	347.4	77.2	694	9	AF237907	AF237907 Homo sapi
6	347.4	77.2	697	9	AF321127	AF321127 Homo sapi
7	340.8	75.7	417	6	AX179785	AX179785 Sequence
8	153.4	34.1	138097	2	AC027787	AC027787 Homo sapi
9	153.4	34.1	138097	2	AC027787	AC027787 Homo sapi
10	153.4	34.1	147788	2	AC015840	AC015840 Homo sapi
11	153.4	34.1	161039	2	AP001327	AP001327 Homo sapi
12	153.4	34.1	166804	2	AP001034	AP001034 Homo sapi
13	153.4	34.1	166804	2	AC090401	AC090401 Homo sapi
14	153.4	29.4	161039	2	AP003127	AP003127 Homo sapi
15	113	25.1	246	6	AX179783	AX179783 Sequence
16	112.4	25.0	166804	2	AP001034	AP001034 Homo sapi
17	69.8	15.5	2805	9	AF367473	AF367473 Homo sapi
18	68.4	15.2	1076	9	AF068288	AF068288 Homo sapi
19	66.8	14.8	720	9	AB022821	AB022821 Homo sapi
20	66.8	14.8	916	9	AB013102	AB013102 Homo sapi
21	66.8	14.8	1619	9	AF237912	AF237912 Homo sapi
22	66.8	14.8	1669	6	AR035695	AR035695 Sequence
23	66.8	14.8	1669	6	AR083580	AR083580 Sequence
24	50.2	11.2	708	6	AX101306	AX101306 Sequence
25	50.2	11.2	1010	10	MUSFCERB	J05019 Mouse mast
26	46	10.2	1628	9	BC008487	BC008487 Homo sapi
27	46	10.2	1646	9	HUMIERB	L35848 Homo sapien
28	46	10.2	1661	6	A68627	A68627 Sequence 3
29	46	10.2	1661	6	AR082014	AR082014 Sequence
30	46	10.2	1661	6	I25809	I25809 Sequence 3
31	44.8	10.0	165139	9	HS436K10	AL031407 Human DNA
32	44.4	9.9	2542	10	RATFCIGA	M22923 Rat Ige Fc
33	44.4	9.9	2545	6	AR123798	AR123798 Sequence 5
34	44.4	9.9	2545	6	I09687	I09687 Sequence 5
35	43.4	9.6	518	10	AF280401	AF280401 Mus muscu
36	43.2	9.6	180017	2	AC073277	AC073277 Homo sapi
37	43.2	9.6	191916	2	AC068760	AC068760 Homo sapi
38	43	9.6	12029	3	AE001379	AE001379 Plasmodiu
39	43	9.6	141540	9	AL138828	AL138828 Human DNA
40	43	9.6	175238	2	AC022654	AC022654 Homo sapi
41	42.4	9.4	11002	5	CHKUPAG	J05188 Chicken uro
42	42.4	9.4	165358	2	AC009020	AC009020 Homo sapi
43	42.4	9.4	186187	9	AC007464	AC007464 Homo sapi
44	42	9.3	61010	9	AL138961	AL138961 Human DNA
45	42	9.3	167830	2	AC027768	AC027768 Homo sapi

ALIGNMENTS

RESULT 1

AX179787

LOCUS AX179787

DEFINITION Sequence 7 from Patent WO0146417.

ACCESSION AX179787

VERSION AX179787.1 GI:15132147

KEYWORDS

SOURCE

ORGANISM human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 450)

Waike, D.W. and Turner, C.A.

Human membrane proteins and polynucleotides encoding the same having homology to cd20 proteins and ige receptors

Patent: WO 0146417-A 7 28-JUN-2001;

JOURNAL

Lexicon Genetics Incorporated (US)

FEATURES

Location/Qualifiers

1..450

/organism="Homo sapiens"

06-AUG-2001


```

DB 273 TTTATATTTCTTTCAGAGATATTCATCTGGGCTCTGTTTGTCTATATTCGTGAGCC 332
QY 301 ttcccaatgcagtgaaagaaacacagaaactctcgggaatttgatcatgag 360
DB 333 TTCCTATATTCAGTGAAGAAACACAGAACTCTGATATATGTAGCCGAATATATG 392
QY 361 acttt 365
DB 393 AATCT 397

RESULT 4
LOCUS AB013103 691 bp mRNA PRI 20-MAR-2001
DEFINITION Homo sapiens mRNA for MS4A5, complete cds.
ACCESSION AB013103
VERSION AB013103.1 GI:11559213
KEYWORDS MS4A5; CD20-like 2.
SOURCE Homo sapiens testis cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Ishibashi, K., Suzuki, M., Sasaki, S. and Imai, M.
AUTHORS Identification of a new multigene four-transmembrane family (MS4A)
TITLE related to CD20, Htm4 and beta subunit of the high-affinity Ige
JOURNAL receptor
MEDLINE Gene 264 (1), 87-93 (2001)
REFERENCE 2 (sites)
AUTHORS Ishibashi, K., Sasaki, S. and Marumo, F.
JOURNAL Cloning of three CD20 homolog from human, putative calcium channels
REFERENCE 3 (bases 1 to 691)
AUTHORS Ishibashi, K.
TITLE Direct Submission
JOURNAL Submitted (20-APR-1998) Kenichi Ishibashi, Tokyo Medical and Dental
UNIVERSITY, 2nd Internal Medicine; Yushima 1-5-45, Bunkyo, Tokyo
113-8519, Japan (E-mail: kishibashi.med2med.tmd.ac.jp,
Tel:81-3-5803-5223, Fax:81-3-5803-0132)
FEATURES
source
location/Qualifiers
1..691
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue.type="testis"
52..654
/gene="MS4A5"
52..654
/gene="MS4A5"
/codon.start=1
/product="MS4A5"
/protein_id="BAB18739.1"
/db_xref="GI:11559214"
/translacion="MDSSTASHPVFLVPPPEITASEVSESTELSATTFSTOSPLQKLF
RKMLIGIOLIEFGIMFSPGVIFLFTLKRPVPPPIFLSGYPMGCVLPINSAGFL
IAYRKTEETLIIISRIIMNLSAIGALIGIILTFGLDNIYICGYSHNSQCAVY
VHFGILITMTFSIETLFIPLPSIILCHSEDDCEQCC"
BASE COUNT 195 a 142 c 117 g 237 t
ORIGIN
Query Match 77.2%; Score 347.4; DB 9; Length 691;
Best Local Similarity 97.0%; Pred. No. 2.5e-67;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 atgagattcaagcagcagacagtcgcgtgttctcgtgatttctccacagaaactcgtct 60
DB 52 ATGGATTCAGACGACGACAGTCGCGGTTCGTGGTATTCCTCCACGAATATCACTGCT 111
QY 61 ttagatatagatccacagaaacttccagcagaccttttaactcaaaagcccttgcaa 120
DB 112 TCAGATATAGATCCACAGAACTTTTCAGCAGACTTTTCAATCAAGCCCTTGCAA 171

```

```

QY 121 aaattatgctagaaaataagaaacttagggactatccagatcctcgttggaaatag 180
DB 172 AATATTTTTCGTAGAAAATGAAAATCTTAGGACATATCCAGATCCTGTTGGAAATATG 231
QY 181 acccttctcttgaggtatcctccttacccttgtaaaccatatacgaagttccc 240
DB 232 ACCCTTCTTTGGAGTATATCTTCCTTTTCAGTTGTAAACCATATCCAGGTTTCCC 291
QY 241 ttatattctcttcagatatccatctcgtggcctcgttttgatcaatcttgagcc 300
DB 292 TTTATATTTCTTTCAGAGATATTCATCTGGGCTCTGTTTGTCTATATTCGTGAGCC 351
QY 301 ttcccaatgcagtgaaagaaacacagaaactctcgggaatttgatcatgag 360
DB 352 TTCCTATATTCAGTGAAGAAACACAGAACTCTGATATATGTAGCCGAATATATG 411
QY 361 acttt 365
DB 412 AATTT 416

RESULT 5
LOCUS AF237907 694 bp mRNA PRI 17-APR-2001
DEFINITION Homo sapiens MS4A5 protein mRNA, complete cds.
ACCESSION AF237907
VERSION AF237907.1 GI:13649400
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 694)
AUTHORS Liang, Y. and Tedder, T.F.
TITLE Identification of a cd20-, fcepsilonibeta-, and htm4-related gene
JOURNAL family: sixteen new ms4a family members expressed in human and
MEDLINE Genomics 72 (2), 119-127 (2001)
REFERENCE 2 (bases 1 to 694)
AUTHORS Liang, Y. and Tedder, T.F.
JOURNAL Direct Submission
JOURNAL Submitted (24-FEB-2000) Immunology, Duke Medical Center, Research
Dr., Durham, NC 27710, USA
FEATURES
source
location/Qualifiers
1..694
/organism="Homo sapiens"
/db_xref="taxon:9606"
55..657
/gene="MS4A5"
55..657
/gene="MS4A5"
/codon.start=1
/product="MS4A5 protein"
/protein_id="AAK37416.1"
/db_xref="GI:13649401"
/translacion="MDSSTASHPVFLVPPPEITASEVSESTELSATTFSTOSPLQKLF
RKMLIGIOLIEFGIMFSPGVIFLFTLKRPVPPPIFLSGYPMGCVLPINSAGFL
IAYRKTEETLIIISRIIMNLSAIGALIGIILTFGLDNIYICGYSHNSQCAVY
VHFGILITMTFSIETLFIPLPSIILCHSEDDCEQCC"
BASE COUNT 198 a 143 c 116 g 237 t
ORIGIN
Query Match 77.2%; Score 347.4; DB 9; Length 694;
Best Local Similarity 97.0%; Pred. No. 2.5e-67;
Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 atgagattcaagcagcagacagtcgcgtgttctcgtgatttctccacagaaactcgtct 60
DB 55 ATGGATTCAGACGACGACAGTCGCGGTTCGTGGTATTCCTCCACGAATATCACTGCT 114
QY 61 ttagatatagatccacagaaacttccagcagaccttttaactcaaaagcccttgcaa 120
DB 115 TCAGATATAGATCCACAGAACTTTTCAGCAGACTTTTCAATCAAGCCCTTGCAA 174

```

QY	121	aaatttttccgtaaaaaaagaataacttaggacataccagaatccgtttggaattatg	180
Db	175	AAATTAATTTGCTGAGAAAATGAAAATCTTAGGACATCTCAGATCTCTGTTGGAAATTATG	234
QY	181	accttctctttggagttatctctcttcttcaaccttglyttaaaccataccaagtctcc	240
Db	235	ACCTTTCTTTTGGAGTTATCTTCTCTTTCACCTTTGTATTAACCATATCCAAAGTTTCC	294
QY	241	tttatattcttcccaagatatacatctctggggtctgtttgttcaattaattctgagcc	300
Db	295	TTTATATTTCTTTACAGATATTCATTCTGGGGCTGTGTTTGTTCATTATTTGTGAGACC	354
QY	301	ttctcaattgcagltgaaaaaacaacccaagaacctggaattttgattacatgatg	360
Db	355	TTTCCTAATTTGCAGTGAAGAAAATACACAGAACTCTGATATATATTGAGCCGAATTAATG	414
QY	361	acctt	365
Db	415	AAATTT	419
RESULT	6		
AF321127			
LOCUS	AF321127	697 bp	mRNA
DEFINITION			PRI
			07-FEB-2001
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
TURES			
SOURCE			
gene			
CDS			
BASE COUNT			
ORIGIN			

QY	1	atgagttcaagaccgcacacagtcgcggtgtgttcctggtatcttcacagaatacagtc	60
Db	57	ATGGATTCACAGACCGCACAGTCGCGGTGTTTCGGATTTCTCCGAGAAATCAGTCT	116
QY	61	tcaagaatagtcacagaaacttcagccacgacactttcaactcaaaagcccttgcaa	120
Db	117	TCAGAAATATGATCCACAGAACTTTCACGCCACACACTTTTCAACTCAAAAGCCCTTGCAA	176
QY	121	aaatatttgcagaaaaatgaaactcttaggactataccagatcccglttggaaattatg	180
Db	177	AAATATTTGCTAGAAAAATGAAAACTTTAGGGACTATCCAGATCCTGTTGGAAATATG	236
QY	181	acccttctcttgaggtatctctcttcacactgtttaaacatatccaaagtttccc	240
Db	237	ACCTTTCTCTTGGAGGTATCTCTCTTTTCACCTTTGTTAAACATATCCAAAGTTTCCC	296
QY	241	ttataattcttcacagatatacattctggygcctcgtgtttgttcataatctcggagcc	300
Db	297	TTTATATTTCTTTTCAGGATATCCATTTCTGGGGCTGTTGTTGTCATTAATTCGGAGCC	356
QY	301	ttcccaattgcagtgaaaaaagaaacccacagaactctcgggaatttggatcaatgatg	360
Db	357	TTCCATATTTGCAAGTGAAGAAAAACACAGAAACTCTGATTAATATTTAGCCGCAATATG	416
QY	361	acttt 365	
Db	417	AATTT 421	
RESULT	7		
AX179785			
LOCUS	AX179785	417 bp	DNA
DEFINITION	Sequence 5 from Patent WO0146417.		PAT
ACCESSION	AX179785		
VERSION	AX179785.1	GI:15132146	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 417)		
JOURNAL	Walke,D.W. and Turner,C.A.		
FEATURES	Human membrane proteins and polynucleotides encoding the same		
Source	having homology to cd20 proteins and Ige receptors		
	Patent: WO 0146417-A-5 28-JUN-2001;		
	Lexicon Genetics Incorporated (US)		
	Location/Qualifiers		
	1..417		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	115 a	91 c	69 g
ORIGIN			142 t
Query Match	75.7%;	Score 340.8;	DB 6;
Best Local Similarity	94.1%;	Pred. No. 7.2e-66;	Length 417;
Matches 354;	Conservative	0;	Mismatches 22;
		Indels	0;
		Gaps	
QY	1	atgattccaagacacgcacacagtcgcggtgttctggtatcttcacagaatacactgct	60
Db	1	ATGGATTCAAGCACCGCACACAGTCGCGTGTTCGTATTTCTCCAGAAATCACTGCT	60
QY	61	tcaaaatagtcacagaaacttcagccacgacactttcaactcaaaagcccttgcaa	120
Db	61	TCAGAAATATGATCCACAGAACTTTCACGCCACACACTTTTCAACTCAAAAGCCCTTGCAA	120
QY	121	aaatatttgcagaaaaatgaaactcttaggactataccagatcccglttggaaattatg	180
Db	121	AAATATTTGCTAGAAAAATGAAAACTTTAGGGACTATCCAGATCCTGTTGGAAATATG	180
QY	181	acccttctcttgaggtatctctcttcacactgtttaaacatatccaaagtttccc	240

Db 181 ACCTTTCTTTGGAGCTATCTCTTTACCTGTTAAACCATATCCAGTTTCCC 240
Qy 241 ttatatcttccagagatcatctcgtggcgtcttcttcttcaattatgagcc 300
Db 241 TTTATATTTCTTTCAGAGATATCCATTCTG6GCTGTTTGTTCATTAATCTTGAGCC 300
Qy 301 ttccatgacagtgaaagaaacacagaacaccccttggaatttgattgattacatgacg 360
Db 301 TTTCCTAATTCGACGTGAAAGAAAACACAGAAACCTGTATCAAACTACATTGTGTT 360
Qy 361 acctcagcattatg 376
Db 361 ATTCTCACCATAATAG 376

RESULT 8
LOCUS AC027787 138097 bp DNA HTG 18-AUG-2000
DEFINITION Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS
SEQUENCE SAMPLING.
AC027787
AC027787.2 GI:9845160
HTG: HTGS_PHA5E0.
KEYWORDS
SOURCE
ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 138097)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
JOURNAL Homo sapiens chromosome 15, clone RP11-196E16
REFERENCE 2 (bases 1 to 138097)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bida, F.,
Boguski, V., Bouckhagalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domingo, M., Doyle, M., Ferreira, P., FitzHugh, M., Gage, D.,
Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,
Melidiri, J., Menus, L., Mihova, T., Miranda, C., Mlenka, Y., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Triggilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L9008
Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
663: contig of 663 bp in length
664 763: gap of 100 bp
764 1464: contig of 701 bp in length
1465 1564: gap of 100 bp
1565 2268: contig of 704 bp in length
2269 2368: gap of 100 bp
2369 3042: contig of 674 bp in length
3043 3142: gap of 100 bp
3143 3620: contig of 678 bp in length
3621 3920: gap of 100 bp
3921 4626: contig of 706 bp in length
4627 4726: gap of 100 bp
4727 5423: contig of 697 bp in length
5424 5523: gap of 100 bp
5524 6234: contig of 711 bp in length
6235 6334: gap of 100 bp
6335 7040: contig of 706 bp in length
7041 7140: gap of 100 bp
7141 7637: contig of 697 bp in length
7638 7937: gap of 100 bp
7938 8632: contig of 695 bp in length
8633 8732: gap of 100 bp
8733 9430: contig of 698 bp in length
9431 9530: gap of 100 bp
9531 10214: contig of 684 bp in length
10215 10314: gap of 100 bp
10315 10980: contig of 666 bp in length
10981 11080: gap of 100 bp
11081 11783: contig of 703 bp in length
11784 11883: gap of 100 bp
11884 12582: contig of 699 bp in length
12583 12682: gap of 100 bp
12683 13389: contig of 707 bp in length
13390 13489: gap of 100 bp
13490 14186: contig of 697 bp in length
14187 14286: gap of 100 bp
14287 14996: contig of 710 bp in length
14997 15096: gap of 100 bp
15097 15796: contig of 700 bp in length
15797 15896: gap of 100 bp
15897 16583: contig of 687 bp in length
16584 16683: gap of 100 bp
16684 17376: contig of 693 bp in length
17377 17476: gap of 100 bp
17477 18178: contig of 702 bp in length
18179 18278: gap of 100 bp
18279 18962: contig of 684 bp in length
18963 19062: gap of 100 bp
19063 19673: contig of 611 bp in length
19674 19773: gap of 100 bp
19774 20436: contig of 663 bp in length
20437 20536: gap of 100 bp
20537 21233: contig of 697 bp in length
21234 21333: gap of 100 bp
21334 22042: contig of 709 bp in length
22043 22142: gap of 100 bp
22143 22808: contig of 666 bp in length
22809 22908: gap of 100 bp
22909 23595: contig of 687 bp in length
23596 23695: gap of 100 bp
23696 24391: contig of 696 bp in length
24392 24491: gap of 100 bp
24492 25190: contig of 699 bp in length
25191 25290: gap of 100 bp
25291 25981: contig of 691 bp in length
25982 26081: gap of 100 bp

```

* 26082 26765: contig of 684 bp in length
* 26766 26865: gap of 100 bp
* 26866 27563: contig of 698 bp in length
* 27564 27663: gap of 100 bp
* 27664 28377: contig of 714 bp in length
* 28378 28477: gap of 100 bp
* 28478 29184: contig of 707 bp in length
* 29185 29284: gap of 100 bp
* 29285 29983: contig of 699 bp in length
* 29984 30083: gap of 100 bp
* 30084 30775: contig of 692 bp in length
* 30776 30875: gap of 100 bp
* 30876 31565: contig of 680 bp in length
* 31566 31665: gap of 100 bp
* 31666 32360: contig of 695 bp in length
* 32361 32460: gap of 100 bp
* 32461 33157: contig of 697 bp in length
* 33158 33257: gap of 100 bp
* 33258 33951: contig of 694 bp in length
* 33952 34051: gap of 100 bp
* 34052 34735: contig of 684 bp in length
* 34736 34835: gap of 100 bp
* 34836 35483: contig of 648 bp in length
* 35484 35583: gap of 100 bp
* 35584 36308: contig of 725 bp in length
* 36309 36408: gap of 100 bp
* 36409 37108: contig of 700 bp in length
* 37109 37208: gap of 100 bp
* 37209 37892: contig of 684 bp in length
* 37893 37992: gap of 100 bp
* 37993 38672: contig of 680 bp in length
* 38673 38772: gap of 100 bp
* 38773 39473: contig of 703 bp in length
* 39476 39575: gap of 100 bp
* 39576 40290: contig of 715 bp in length
* 40291 40390: gap of 100 bp
* 40391 41087: contig of 697 bp in length
* 41088 41187: gap of 100 bp
* 41188 41891: contig of 704 bp in length
* 41892 41991: gap of 100 bp
* 41992 42679: contig of 688 bp in length
* 42680 42779: gap of 100 bp
* 42780 43500: contig of 721 bp in length
* 43501 43600: gap of 100 bp
* 43601 44295: contig of 695 bp in length
* 44296 44395: gap of 100 bp
* 44396 45088: contig of 693 bp in length
* 45089 45188: gap of 100 bp
* 45189 45880: contig of 692 bp in length
* 45881 45980: gap of 100 bp
* 45981 46683: contig of 703 bp in length
* 46684 46783: gap of 100 bp
* 46784 47481: contig of 698 bp in length
* 47482 47581: gap of 100 bp
* 47582 48290: contig of 709 bp in length
* 48291 48390: gap of 100 bp
* 48391 49077: contig of 687 bp in length
* 49078 49177: gap of 100 bp
* 49178 49879: contig of 702 bp in length
* 49880 49979: gap of 100 bp
* 49980 50684: contig of 705 bp in length
* 50685 50784: gap of 100 bp
* 50785 51507: contig of 723 bp in length
* 51508 51607: gap of 100 bp
* 51608 52320: contig of 713 bp in length
* 52321 52420: gap of 100 bp
* 52421 53133: contig of 713 bp in length
* 53134 53233: gap of 100 bp
* 53234 535933: contig of 700 bp in length
* 53934 54033: gap of 100 bp
* 54034 54703: contig of 670 bp in length

```

Query Match

34.1%; Score 153.4; DB 2; Length 138097;

Best Local Similarity 96.3%; Pred. No. 2.8e-24;
Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 atgagttcaagcaccgacgacgtcgtgtgttctgtatctcccccgaatactgct 60

DB 109596 ATGGATTCAGACCGCACACAGTCGGGTTCGTATTCCTCCAGAACTACTCT 109655

QY 61 tcagataatgacacagaaacttcacgacacacttcaactcaagaacccctgcaa 120

DB 109656 TCAGATATNGATCCACAGAACTTTCAGCACAGACTTTTCACTCAAAAGCCCTTGC 109715

QY 121 aaattatttgcctagaaaatgaactcttaggactatccaga 163

DB 109716 AAATTATTTGCTAGAAAATGAAATCTTAGCGGTAGTAAGA 109758

RESULT 9 AC027787/c 138097 bp DNA HTG 18-AUG-2000

LOCUS Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS

DEFINITION SEQUENCE SAMPLING.

ACCESSION AC027787

VERSION AC027787.2 GI:9845160

KEYWORDS HTG; HTGS; PHASE0.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 138097)

AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.

TITLE Homo sapiens chromosome 15, clone RP11-196E16

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 138097)

AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Boguslavsky, L., Bouhassira, B., Brown, A., Burt, G.,

Campopiano, A., Castle, A., Choe, Y., Collinge, M., Collins, S.,

Collins, S., Dominko, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,

Galagan, J., Gardina, S., Ginde, S., Goyle, M., Graham, L.,

Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,

Howland, J., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,

Klein, J., Lacombe, K., Lamazares, R., Landers, T., Lehotzky, J.,

Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,

McCarthy, M., McEwan, P., McKernan, K., McPherson, R.,

Meldrum, J., Meneses, L., Mihov, T., Miranda, C., Mlepe, V., Morrow, J.,

Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

O'Neil, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,

Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rotman, D.,

Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,

Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

Teafaye, S., Theodore, J., Tirrell, A., Travers, M., Trigg, J.,

Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,

Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: U908
Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads

```
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
1
* 663: contig of 663 bp in length
* 664 763: gap of 100 bp
* 764 1464: contig of 701 bp in length
* 1465 1564: gap of 100 bp
* 1565 2268: contig of 704 bp in length
* 2269 2368: gap of 100 bp
* 2369 3042: contig of 674 bp in length
* 3043 3142: gap of 100 bp
* 3143 3820: contig of 678 bp in length
* 3821 3920: gap of 100 bp
* 3921 4626: contig of 706 bp in length
* 4627 4726: gap of 100 bp
* 4727 5423: contig of 697 bp in length
* 5424 5523: gap of 100 bp
* 5524 6234: contig of 711 bp in length
* 6235 6334: gap of 100 bp
* 6335 7040: contig of 706 bp in length
* 7041 7140: gap of 100 bp
* 7141 7837: contig of 697 bp in length
* 7838 7937: gap of 100 bp
* 7938 8632: contig of 695 bp in length
* 8633 8732: gap of 100 bp
* 8733 9430: contig of 698 bp in length
* 9431 9530: gap of 100 bp
* 9531 10214: contig of 684 bp in length
* 10215 10314: gap of 100 bp
* 10315 10980: contig of 666 bp in length
* 10981 11080: gap of 100 bp
* 11081 11783: contig of 703 bp in length
* 11784 11883: gap of 100 bp
* 11884 12582: contig of 699 bp in length
* 12583 12682: gap of 100 bp
* 12683 13389: contig of 707 bp in length
* 13390 13489: gap of 100 bp
* 13490 14186: contig of 697 bp in length
* 14187 14286: gap of 100 bp
* 14287 14996: contig of 710 bp in length
* 14997 15096: gap of 100 bp
* 15097 15796: contig of 700 bp in length
* 15797 15896: gap of 100 bp
* 15897 16583: contig of 687 bp in length
* 16584 16683: gap of 100 bp
* 16684 17376: contig of 693 bp in length
* 17377 17476: gap of 100 bp
* 17477 18178: contig of 702 bp in length
* 18179 18278: gap of 100 bp
* 18279 18962: contig of 684 bp in length
* 18963 19062: gap of 100 bp
* 19063 19673: contig of 611 bp in length
* 19674 19773: gap of 100 bp
* 19774 20436: contig of 663 bp in length
* 20437 20536: gap of 100 bp
* 20537 21233: contig of 697 bp in length
* 21234 21333: gap of 100 bp
* 21334 22042: contig of 709 bp in length
* 22043 22142: gap of 100 bp
* 22143 22808: contig of 666 bp in length
* 22809 22908: gap of 100 bp
* 22909 23595: contig of 687 bp in length
* 23596 23695: gap of 100 bp
* 23696 24391: contig of 696 bp in length
* 24392 24491: gap of 100 bp
* 24492 25190: contig of 699 bp in length
* 25191 25290: gap of 100 bp
* 25291 25981: contig of 691 bp in length
25982 26081: gap of 100 bp
26082 26765: contig of 664 bp in length
26766 26865: gap of 100 bp
26866 27563: contig of 698 bp in length
27564 27663: gap of 100 bp
27664 28377: contig of 714 bp in length
28378 28477: gap of 100 bp
28478 29184: contig of 707 bp in length
29185 29284: gap of 100 bp
29285 29983: contig of 699 bp in length
29984 30083: gap of 100 bp
30084 30775: contig of 692 bp in length
30776 30875: gap of 100 bp
30875 31565: contig of 690 bp in length
31566 31665: gap of 100 bp
31666 32360: contig of 695 bp in length
32361 32460: gap of 100 bp
32461 33157: contig of 697 bp in length
33158 33257: gap of 100 bp
33258 33951: contig of 694 bp in length
33952 34051: gap of 100 bp
34052 34735: contig of 684 bp in length
34736 34835: gap of 100 bp
34836 35483: contig of 648 bp in length
35484 35583: gap of 100 bp
35584 36308: contig of 725 bp in length
36309 36408: gap of 100 bp
36409 37108: contig of 700 bp in length
37109 37208: gap of 100 bp
37209 37892: contig of 684 bp in length
37893 37992: gap of 100 bp
37993 38672: contig of 680 bp in length
38673 38772: gap of 100 bp
38773 39473: contig of 703 bp in length
39476 39575: gap of 100 bp
39576 40290: contig of 715 bp in length
40291 40390: gap of 100 bp
40391 41087: contig of 697 bp in length
41088 41187: gap of 100 bp
41188 41891: contig of 704 bp in length
41892 41991: gap of 100 bp
41992 42679: contig of 688 bp in length
42680 42779: gap of 100 bp
42780 43500: contig of 721 bp in length
43501 43600: gap of 100 bp
43601 44295: contig of 695 bp in length
44296 44395: gap of 100 bp
44396 45088: contig of 693 bp in length
45089 45188: gap of 100 bp
45189 45880: contig of 692 bp in length
45881 45980: gap of 100 bp
45981 46683: contig of 703 bp in length
46684 46783: gap of 100 bp
46784 47481: contig of 698 bp in length
47482 47581: gap of 100 bp
47582 48290: contig of 709 bp in length
48291 48390: gap of 100 bp
48391 49077: contig of 687 bp in length
49078 49177: gap of 100 bp
49178 49879: contig of 702 bp in length
49880 49979: gap of 100 bp
49980 50684: contig of 705 bp in length
50685 50784: gap of 100 bp
50785 51507: contig of 723 bp in length
51508 51607: gap of 100 bp
51608 52320: contig of 713 bp in length
52321 52420: gap of 100 bp
52421 53133: contig of 713 bp in length
53134 53233: gap of 100 bp
53234 53933: contig of 700 bp in length
53934 54033: gap of 100 bp
54034 54703: contig of 670 bp in length
```


REFERENCE	TITLE	JOURNAL	AUTHORS
1	Hattori M., Ishii K., Toyoda A., Taylor T.D., Hong-Seog P., Fujiyama A., Yada T., Totoki Y., Watanabe H. and Sakaki Y.	Homo sapiens 161,039 genomic DNA of 11q	
2	(bases 1 to 161039)		
3	Hattori M., Ishii K., Toyoda A., Taylor T.D., Hong-Seog P., Fujiyama A., Yada T., Totoki Y., Watanabe H. and Sakaki Y.	Direct Submission	
4	Submitted (26-JUN-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)		
5	Genome Center		
6	Center: RIKEN Genomic Sciences Center(GSC)		
7	Center code: RIKEN		
8	Web site: http://hgp.gsc.riken.go.jp/		
9	Contact: hattori@gsc.riken.go.jp		
10	Project information		
11	Center project name: Humpraf11		
12	Center clone name: RP11-71066		
13	Summary Statistics		
14	Sequencing vector: PCR products, 100% of reads		
15	Chemistry: Dye-terminator ET-amersham; 100% of reads		
16	Assembly program: Phrap; version 0.990329		
17	Consensus quality: 157456 bases at least Q40		
18	Consensus quality: 158960 bases at least Q30		
19	Consensus quality: 159457 bases at least Q20		
20	Insert size: 159639; sum-of-contigs		
21	Quality coverage: 8.61x in Q20 bases; sum-of-contigs		
22	NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.		
23	1	42322 contig of 42322 bp in length	
24	42423	63418 contig of 20996 bp in length	
25	63519	77969 contig of 14451 bp in length	
26	78070	93182 contig of 15113 bp in length	
27	93283	104093 contig of 10811 bp in length	
28	104194	115702 contig of 11550 bp in length	
29	115803	125183 contig of 9381 bp in length	
30	125284	132558 contig of 7251 bp in length	
31	132659	137219 contig of 4561 bp in length	
32	137320	143276 contig of 5566 bp in length	
33	143377	148922 contig of 5566 bp in length	
34	149023	154159 contig of 5137 bp in length	
35	154260	157686 contig of 3427 bp in length	
36	157787	159234 contig of 1448 bp in length	
37	159335	161039 contig of 1705 bp in length.	
38	NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.		
39	1	42322: contig of 42322 bp in length	
40	42323	42423: gap of 100 bp	
41	42423	63418: contig of 20996 bp in length	
42	63419	63519: gap of 100 bp	
43	63519	77969: contig of 14451 bp in length	
44	77970	78069: gap of 100 bp	
45	78070	93182: contig of 15113 bp in length	
46	93183	93282: gap of 100 bp	
47	93283	104093: contig of 10811 bp in length	
48	104094	104193: gap of 100 bp	
49	104194	115702: contig of 11550 bp in length	
50	115703	115802: gap of 100 bp	

Query Match	34.1%	Score 153.4	DB 2	Length 161039
Best Local Similarity	96.3%	Pred. No. 2.8e-24		
Matches 157	Conservative 0	Mismatches 6	Indels 0	Gaps 0
QY 1	atgagttcaagaccgcgcacacagtcgcggtgttcctggtattctctccgaataactgct	60		
DB 149618	ATGGATATCAAGACCGCACACAGTCGCGGTTCGTGGATTTCTCCAGAAATCACTGCT	149559		
DB 149558	TCAGAAATATGAGTCCACAGAACTTCACGACGACACTTTCAATCAATCAAGCCCTTGCAA	149456		
QY 121	aaattattgtctagaataatgaataatctaggagcattaccaga	163		
DB 149496	AAATTAATTTGCTAGAAAATGAAAATCTTAGGCGGTAGTAAGA	149456		

RESULT 12
AP001034/c

LOCUS
DEFINITION
Homo sapiens chromosome 11 clone RP11-729B4 map 11q12, WORKING DRAFT SEQUENCE, 19 unordered pieces.
AP001034

ACCESSION
AP001034.4 GI:11176992

VERSION
HTG: HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS
Homo sapiens DNA, clone:RP11-729B4.

SOURCE
Homo sapiens

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 166804)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission

TITLE
JOURNAL
Submitted (06-JAN-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0043, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
On Nov 14, 2000 this sequence version replaced gi:1117704.

COMMENT
Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
Project Information
Center project name: Humdraft11
Center clone name: RP11-729B4
----- Summary Statistics -----
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator; PCR-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 162254 bases at least Q40
Consensus quality: 163858 bases at least Q30
Consensus quality: 164544 bases at least Q20
Insert size: 165004; sum-of-contigs
Quality coverage: 9.28x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 19 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved
1
22398 contig of 22398 bp in length
22499 45238 contig of 22740 bp in length
45339 64278 contig of 18940 bp in length
64379 77579 contig of 13201 bp in length
77680 93718 contig of 16039 bp in length
93819 108057 contig of 14229 bp in length
108158 117006 contig of 8849 bp in length
117107 126302 contig of 9196 bp in length
126403 133479 contig of 7077 bp in length
133580 144025 contig of 6846 bp in length
140526 144653 contig of 4128 bp in length
144754 149354 contig of 4601 bp in length
149455 154522 contig of 5068 bp in length
154623 155919 contig of 1997 bp in length
156020 159504 contig of 3485 bp in length
159605 159679 contig of 75 bp in length
159780 162418 contig of 2639 bp in length
162519 165383 contig of 2665 bp in length
165484 166804 contig of 1321 bp in length.

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 22398: contig of 22398 bp in length
* 22399 22498: gap of 100 bp
* 22499 45238: contig of 22740 bp in length
* 45239 45338: gap of 100 bp
* 45339 64278: contig of 18940 bp in length
* 64279 64378: gap of 100 bp
* 64379 77579: contig of 13201 bp in length
* 77580 77679: gap of 100 bp
* 77680 93718: contig of 16039 bp in length
* 93719 93818: gap of 100 bp
* 93819 108057: contig of 14239 bp in length
* 108058 108157: gap of 100 bp
* 108158 117006: contig of 8849 bp in length
* 117007 117106: gap of 100 bp
* 117107 126302: contig of 9196 bp in length
* 126303 126402: gap of 100 bp
* 126403 1333479: contig of 7077 bp in length
* 133480 133579: gap of 100 bp
* 133580 140425: contig of 6846 bp in length
* 140426 140525: gap of 100 bp
* 140526 144653: contig of 4128 bp in length
* 144654 144753: gap of 100 bp
* 144754 349354: contig of 4601 bp in length
* 349355 149454: gap of 100 bp
* 149455 154522: contig of 5068 bp in length
* 154523 154622: gap of 100 bp
* 154623 155919: contig of 1297 bp in length
* 155920 156019: gap of 100 bp
* 156020 159504: contig of 3485 bp in length
* 159505 159604: gap of 100 bp
* 159605 159679: contig of 75 bp in length
* 159680 159779: gap of 100 bp
* 159780 162418: contig of 2639 bp in length
* 162419 162518: gap of 100 bp
* 162519 165383: contig of 2865 bp in length
* 165384 165483: gap of 100 bp
* 165484 166804: contig of 1321 bp in length.
*
* FEATURES
* source
* 1. 166804
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="11"
* /map="11q12"
* /clone="RP11-729B4"
*
* misc.feature
* 1. 22398
* /note="assembly_fragment"
*
* misc.feature
* 22499. 45238
* /note="assembly_fragment"
*
* misc.feature
* 45339. 64278
* /note="assembly_fragment"
*
* misc.feature
* 64379. 77579
* /note="assembly_fragment"
*
* misc.feature
* 77680. 93718
* /note="assembly_fragment"
*
* misc.feature
* 93819. 108057
* /note="assembly_fragment"
*
* misc.feature
* 108158. 117006
* /note="assembly_fragment"
*
* misc.feature
* 117107. 126302
* /note="assembly_fragment"
*
* misc.feature
* 126403. 133479
* /note="assembly_fragment"
*
* misc.feature
* 133580. 140425
* /note="assembly_fragment"
*
* misc.feature
* 140526. 144653

```



```

misc_feature /note="assembly_fragment"
144754 .149354
misc_feature /note="assembly_fragment"
149455 .154322
misc_feature /note="assembly_fragment"
154623 .155919
misc_feature /note="assembly_fragment clone_end:SP6 vector_side:right"
156020 .159504
misc_feature /note="assembly_fragment"
159605 .159679
misc_feature /note="assembly_fragment clone_end:T7 vector_side:left"
159780 .162418
misc_feature /note="assembly_fragment"
162519 .165383
misc_feature /note="assembly_fragment"
165484 .166804
misc_feature /note="assembly_fragment"
165484 .166804
BASE COUNT 50706 a 31523 c 32296 g 50478 t 1801 others
ORIGIN

```

```

Query Match 34.1%; Score 153.4; DB 2; Length 166804;
Best Local Similarity 96.3%; Pred. No. 2.8e-24;
Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 1 atgagttcaagcagcagcagcagtcggtgttctgtatcttcctccagaatacactgct 60
DB 11755 ATGGATTCAAGCAGCCGACACGACGTCGGGTGTTTCTGTTTCTTCCAGAAATCACTGCT 11696
OY 61 taagaatagtatccacagaactttcagccagcagcctttcaactcaagcccttgcaa 120
DB 11695 TCAGATTATGAGTCCACAGAACTTTCAGCCAGCAGCCTTTCAACTCAAGCCCTTGCAA 11636
OY 121 aaattattgtctagaaaaatgaataatcttagggactatccaga 163
DB 11635 AAATTATTTGCTAGAAAAATGAATACTTAGGGGTAAGTACA 11593

```

```

RESULT 13
AC090401/C 166820 bp DNA HTG 21-JUL-2001
LOCUS Homo sapiens chromosome 11 clone RP11-729B4 map 11, WORKING DRAFT
DEFINITION
SEQUENCE, 2 ordered pieces.
AC090401
AC090401.2 GI:14971278
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTROP.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 166820)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
JOURNAL Homo sapiens chromosome 11, clone RP11-729B4
TITLE Unpublished

```

```

REFERENCE 1 (bases 1 to 166820)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barra, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
Camara, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardina, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagges, B., Heaford, A., Horton, L., Hulme, W., Illay, I., Johnson, R.,
Jones, C., Karatas, A., LaRoque, K., Lamazares, R., Landers, T.,
Lehoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrum, J., Menus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retter, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P.,
Soune, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,

```

```

TITLE
JOURNAL
COMMENT

```

Travers, M., Travis, N., Trigilio, J., Vassiliou, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 20, 2001 this sequence version replaced gi:12958037.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 729_B_4
Center clone name: 729_B_4

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 166378 bases at least Q40
Consensus quality: 166639 bases at least Q30
Consensus quality: 166692 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 166720; sum-of-contigs
Quality coverage: 12.2 in Q20 bases; agarose-fp
Quality coverage: 12.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 1096 1195: contig of 1095 bp in length
* 1196 166820: contig of 165625 bp in length.
Location/Qualifiers
1. 166820
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-729B4"
/clone_id="RP11-729B4 Human Male BAC"
1. 1095
/note="assembly_fragment"
1196. 166820
/note="assembly_fragment"
clone_end:T7
vector_side:right"

```

FEATURES
source
1. 166820
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-729B4"
/clone_id="RP11-729B4 Human Male BAC"
1. 1095
/note="assembly_fragment"
1196. 166820
/note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT 46436 a 31826 c 32895 g 55563 t 100 others
ORIGIN

```

```

Query Match 34.1%; Score 153.4; DB 2; Length 166820;
Best Local Similarity 96.3%; Pred. No. 2.8e-24;
Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 1 atgagttcaagcagcagcagcagtcggtgttctgtatcttcctccagaatacactgct 60
DB 144100 ATGGATTCAAGCAGCCGACACGACGTCGGGTGTTTCTGTTTCTTCCAGAAATCACTGCT 144041
OY 61 taagaatagtatccacagaactttcagccagcagcctttcaactcaagcccttgcaa 120
DB 144040 TCAGATTATGAGTCCACAGAACTTTCAGCCAGCAGCCTTTCAACTCAAGCCCTTGCAA 143981
OY 121 aaattattgtctagaaaaatgaataatcttagggactatccaga 163

```

DB 143980 AATATTGCTAGAAAATGAAATCTTAGGGTAGTAGA 143938

RESULT 14

AP003127

LOCUS Homo sapiens chromosome 11 clone RP11-71066 map 11q, WORKING DRAFT

DEFINITION SEQUENCE, 15 unordered pieces.

ACCESSION

AP003127

AP003127.1 GI:12597183

KEYWORDS HTGS, PHASE1, HTGS, DRAFT.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 42322: contig of 42322 bp in length
42323 42422: gap of 100 bp
42423 63418: contig of 20966 bp in length
63419 63518: gap of 100 bp
63519 77969: contig of 14451 bp in length
77970 78069: gap of 100 bp
78070 93182: contig of 15113 bp in length
93183 93282: gap of 100 bp
93283 104093: contig of 10811 bp in length
104094 104193: gap of 100 bp
104194 115702: contig of 11509 bp in length
115703 115802: gap of 100 bp
115803 125183: contig of 9981 bp in length
125184 125283: gap of 100 bp
125284 132558: contig of 7275 bp in length
132559 132658: gap of 100 bp
132659 137219: contig of 4561 bp in length
137220 137319: gap of 100 bp
137320 143276: contig of 5957 bp in length
143277 143376: gap of 100 bp
143377 148922: contig of 5546 bp in length
148923 149022: gap of 100 bp
149023 154159: contig of 5137 bp in length
154160 154259: gap of 100 bp
154260 157686: contig of 3427 bp in length
157687 157786: gap of 100 bp
157787 159234: contig of 1448 bp in length
159235 159335: gap of 100 bp
159335 161039: contig of 1705 bp in length.

FEATURES

SOURCE

1. 161039

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q"

/clone="RP11-71066"

1. 42322

/note="assembly_fragment"

42423. 63418

/note="assembly_fragment"

63519. 77969

/note="assembly_fragment"

78070. 93182

/note="assembly_fragment"

93283. 104093

/note="assembly_fragment"

104194. 115702

/note="assembly_fragment"

115803. 125183

/note="assembly_fragment"

125284. 132558

/note="assembly_fragment"

132659. 137219

/note="assembly_fragment"

137320. 143276

/note="assembly_fragment"

143377. 148922

/note="assembly_fragment"

149023. 154159

/note="assembly_fragment"

154260. 157686

/note="assembly_fragment"

157787. 159234

/note="assembly_fragment"

159335. 161039

/note="assembly_fragment"

/note="assembly_fragment"

BASE COUNT 49155 a 31166 c 31523 g 47795 t 1400 others

* NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces

* is not known and their order in this sequence record is

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
Project Information
Center project name: Humdrill1
Center clone name: RP11-71066
Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 157456 bases at least Q40
Consensus quality: 158960 bases at least Q30
Consensus quality: 159457 bases at least Q20
Insert size: 159639; sum-of-contigs
Quality coverage: 8.61x in Q20 bases; sum-of-contigs

Query Match 29.4%; Score 132.4; DB 2; Length 161039;
Best Local Similarity 95.8%; Pred. No. 1.2e-19;
Matches 136; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 148 ttagggactatccagatcctgttggaaatagacctttcttggagttatcttcctt 207
|||
Db 42956 TTCAGACTATCCAGATCCGTTGGAATATGACCTTTCTTTGAGTTATCTTCCTT 43015

QY 208 ttcaacttgttaaacacatccaaagttcccttataatttcttcagatatccattc 267
|||
Db 43016 TTCACCTGTAAACCATATCCAGGTTCCCTTATATTCTTTCAGGATATCCATTC 43075

QY 268 ttgggctctgtttgtcat 289
|||
Db 43076 TGGGCTCTGTTTGGTGAGTA 43097

RESULT 15

ACUS 119783

DEFINITION AX179783 246 bp DNA PAT 06-AUG-2001

SEQUENCE 3 from Patent WO0146417.

ACCESSION AX179783

VERSION AX179783.1 GI:15132145

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 246)

AUTHORS Walke,D.W. and Turner,C.A.

TITLE Human membrane proteins and polynucleotides encoding the same

JOURNAL Patent: WO 0146417-A 3 28-JUN-2001;

Lexicon Genetics Incorporated (US)

FEATURES

source location/Qualifiers

1..246

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 60 a 45 c 47 g 94 t

ORIGIN

Query Match 25.1%; Score 113; DB 6; Length 246;

Best Local Similarity 100.0%; Pred. No. 2e-15;

Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 338 ttgggaatttgatcatgatgacttgcagcatatgaattatcatctctgcctt 397
|||
Db 134 TGGGAATTTGATACATGATGACCTTCAGCATTTGATATTCATTCTCTGCTT 193

QY 398 tctcaatttgggggagccactcagagagattgtgataaactgttctga 450
|||
Db 194 TCTCAATTTGGGGTGCACACAGAGATTTGATGTGAACAATGTGTGTA 246

Search completed: November 28, 2001, 23:48:11
Job time: 8134 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2001, 23:12:03 ; Search time 204.6 Seconds

(without alignments)
1885.613 Million cell updates/sec

Title: US-09-735-712-7

Perfect score: 450

Sequence: 1 atgattcaagcaccgcgcaca.....attgtgacaattgttga 450

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

tal number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	450	100.0	450	22	Novel human membra
2	347.4	77.2	603	22	Novel human membra
3	347.4	77.2	676	22	DNA sequence encod
4	347.4	77.2	747	22	Human secreted pro
5	340.8	75.7	417	22	Novel human membra
6	280.8	62.4	372	21	Human secreted pro
7	234	52.0	468	22	Human secreted pro
8	113	25.1	246	22	Novel human membra
9	79	17.6	1563	21	cDNA encoding a bo
10	73.6	16.4	936	22	Oligonucleotide D1
11	73.6	16.4	936	22	Oligonucleotide D1

12	73.6	16.4	936	22	AAFS8257
13	73.6	16.4	936	22	AAFS8259
14	73.6	16.4	936	22	AAFS8262
15	73.6	16.4	936	22	AAFS8255
16	71.4	15.9	936	22	AAFS8252
17	71.4	15.9	936	22	AAFS8254
18	71.4	15.9	936	22	AAFS8257
19	71.4	15.9	936	22	AAFS8259
20	71.4	15.9	936	22	AAFS8262
21	71.4	15.9	936	22	AAFS8255
22	66.8	14.8	1330	21	AAAI6693
23	66.8	14.8	1669	20	AAAI15104
24	66.8	14.8	1670	20	AAAI15104
25	50.2	11.2	708	22	AAAI15104
26	48.6	10.8	708	22	AAAI15104
27	46	10.2	1661	17	AAAI15104
28	46	10.2	1661	19	AAAI15104
29	44.4	9.9	244	22	AAAI15104
30	44.4	9.9	2545	11	AAAI15104
31	44.4	9.9	2545	12	AAAI15104
32	44.4	9.9	2545	14	AAAI15104
33	43.6	9.7	244	22	AAAI15104
34	43	9.6	6888	21	AAAI15104
35	42	9.3	1474	13	AAAI15104
36	40.6	9.0	1060	20	AAAI15104
37	40.4	9.0	1474	17	AAAI15104
38	40.4	9.0	1474	17	AAAI15104
39	40.4	9.0	1474	21	AAAI15104
40	40.4	9.0	1474	22	AAAI15104
41	40.4	9.0	1476	10	AAAI15104
42	40.4	9.0	1476	13	AAAI15104
43	40.4	9.0	1476	17	AAAI15104
44	40.4	9.0	1476	19	AAAI15104
45	40.4	9.0	1476	20	AAAI15104

ALIGNMENTS

RESULT 1	
AA504282	standard; CDNA; 450 BP.
ID	AA504282
XX	AA504282;
AC	26-SEP-2001 (first entry)
DT	
XX	Novel human membrane protein #4 cDNA sequence.
DE	
XX	Human; membrane protein; membrane receptor; Ige receptor; CD20;
KW	physiological disorder; ss.
KW	
XX	Homo sapiens.
OS	
XX	
FT	Key
FT	1.450
FT	CD5
FT	/tag= a
FT	/product= "Human membrane protein #4"
PN	WO200146417-A2.
PD	28-JUN-2001.
XX	
PF	12-DEC-2000; 2000MO-US33742.
XX	
PR	22-DEC-1999; 9905-0171567.
XX	
PA	(LEXI-) LEXICON GENETICS INC.
XX	
PI	Walke DW, Turner CA.
XX	
DR	WPI, 2001-408646/43.
DR	P-PSDB; AAU01213.

XX Polynucleotide encoding novel human membrane protein, useful for
PT identifying agonist, antagonist or modifiers or for producing
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
PT applications
XX
XX
PS Disclosure; Page 31; 32pp; English.
XX
XX The present sequence encodes for novel human membrane protein #4.
CC Human membrane protein #4 is 1 of 4 human membrane proteins
CC (AAU01210-AAU01213) given in the present invention. These membrane
CC proteins share structural similarity with membrane receptors such as
CC the IGF receptor and mammalian CD20. The novel human membrane proteins
CC are useful for identifying agonists, antagonists and modulators of the
CC membrane proteins, and for producing antibodies specific to the
CC membrane proteins. The membrane proteins can be used for diagnosis,
CC drug screening, pharmacogenomic applications, clinical trial monitoring
CC and the treatment of physiologic disorders and diseases. The
CC polynucleotides encoding the membrane proteins can be used to generate
CC PCR primers or probes to identify mutations associated with a particular
CC disease.

Sequence 450 BP; 120 A; 94 C; 77 G; 159 T; 0 other;

Query Match	100.0%;	Score 450;	DB 22;	Length 450;
Best Local Similarity	100.0%;	Pred. No. 1.3e-110;		
Matches 450; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0

QY	1	atgattccaagcaacgcagcaacagctcggtttcttggtattcttcctccagaataacatgct	60
Db	1	atgattccaagcaacgcagcaacagctcggtttcttggtattcttcctccagaataacatgct	60
QY	61	tcgaatacatgcttcacagaacatttcagccacgacctttcaactcaataagcccttgcga	120
Db	61	tcgaatacatgcttcacagaacatttcagccacgacctttcaactcaataagcccttgcga	120
QY	121	aaattatttctaaaaaaatgaatatcttggagactatccagatccctgtttggaattatg	180
Db	121	aaattatttctaaaaaaatgaatatcttggagactatccagatccctgtttggaattatg	180
QY	181	acctttctcttggagatcatcttcctcttcaccttgtttaacacatatccaaagtttccc	240
Db	181	acctttctcttggagatcatcttcctcttcaccttgtttaacacatatccaaagtttccc	240
QY	241	tttatatttcttcacagaatcatcttcctgggctcgttttgttcaatcaattcttggagcc	300
Db	241	tttatatttcttcacagaatcatcttcctgggctcgttttgttcaatcaattcttggagcc	300
QY	301	ttcccaattcgagtgaaaaaagaaacacacagaacctcttggaatttggattacatgatg	360
Db	301	ttcccaattcgagtgaaaaaagaaacacacagaacctcttggaatttggattacatgatg	360
QY	361	accttcaagcaatttggaattatccaattcttcgccttcccaatttggggtgccacta	420
Db	361	accttcaagcaatttggaattatccaattcttcgccttcccaatttggggtgccacta	420
QY	421	gaggaatcgatattgtaacaaatgtgttga	450
Db	421	gaggaatcgatattgtaacaaatgtgttga	450

RESULT	2
AAS04279	
ID	AAS04279 standard; cDNA; 603 BP

AC AAS04279;

DT 26-SEP-2001 (first entry)
 XX

Novel human membrane protein #1 cDNA sequence

KW Human; membrane protein; membrane receptor; IgE receptor; CD20;

KW	physiological disorder; ss.
XX	
OS	
XX	Homo sapiens.
XX	
FH	Key
FT	1..603
FT	location/Qualifiers
FT	/taag= a
FT	/product= "Human membrane protein #1"
FT	/cran1_except= (pos:595..600,aa:cys)

PN WO200146417-A2.
YY

PD 28-JUN-2001
YY

PF 12-DEC-2000; 2000WO-US33742.
YY

PR 22-DEC-1999; 99US-0171567.
 VY

PA (LEXI-) LEXICON GENETICS INC.
XX

PI walke DW, Turner CA;
VY

DR WPI; 2001-408646/43.
DR R-DEDDA; 2001-408646/43.

[illegible]

identifying agonist, antagonist or modifiers or for producing antibodies useful in therapeutic and pharmaceutical applications.

PT applications -
XX

PS Claim 1; Page 29; 32pp; English.
XX

The present sequence encodes for novel human membrane protein #1. Human membrane protein #1 is 1 of 4 human membrane proteins (AAU01210-AAU01213) given in the present invention. These membrane proteins share structural similarity with membrane receptors such as the IGE receptor and mammalian CD20. The novel human membrane proteins are useful for identifying agonists, antagonists and modulators of the membrane proteins, and for producing antibodies specific to the membrane proteins. The membrane proteins can be used for diagnosis, drug screening, pharmacogenomic applications, clinical trial monitoring and the treatment of physiological disorders and diseases. The polynucleotides encoding the membrane proteins can be used to generate PCR primers or probes to identify mutations associated with a particular disease.

Sequence 603 BP; 162 A; 125 C; 104 G; 212 T; 0 other;

Query Match	77.2%;	Score 347.4;	DB 22;	Length 603;
Best Local Similarity	97.0%;	Pred. No. 2.8e-83;		
Matches 354; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0;

QY	1	atgagatcaagacgcgcacacagtcggtcttcgtgatctccccaagaatcaagct	60
Db	1	atgagatcaagacgcgcacacagtcggtcttcgtgatctccccaagaatcaagct	60
QY	1	atgagatcaagacgcgcacacagtcggtcttcgtgatctccccaagaatcaagct	60
Db	1	atgagatcaagacgcgcacacagtcggtcttcgtgatctccccaagaatcaagct	60
QY	61	tcaagaatctgagtcacacagaacttcaagcagacctttcaactcaaaagcccttgcaa	120
Db	61	tcaagaatctgagtcacacagaacttcaagcagacctttcaactcaaaagcccttgcaa	120
QY	121	aaatattctgcagagaaaatgaaaattcttgagacatcccaatccctgtttggaattag	180
Db	121	aaatattctgcagagaaaatgaaaattcttgagacatcccaatccctgtttggaattag	180
QY	181	accttctccttggagatcatcttccttcaactgtttaaaccataccaagaattcc	240
Db	181	accttctccttggagatcatcttccttcaactgtttaaaccataccaagaattcc	240
QY	241	tttatattcttctgaagatatacatctcgtgggtcgtttgttatataatctcggagcc	3000
Db	241	tttatattcttctgaagatatacatctcgtgggtcgtttgttatataatctcggagcc	3000

Db 93 tcagaatatgtccacagaacttccagccagcacttccaactcaagcccttgc aa 152

CC determine which patients may be in need of restorative therapy.

CC The GENSET polypeptides may be used as antigens in the production of
 CC antibodies and in assays to identify modulators (agonists and
 CC antagonists) of GENSET polypeptide expression and activity. The
 CC present sequence is a GENSET nucleic acid of the invention.
 CC
 SQ Sequence 747 BP; 240 A; 145 C; 121 G; 241 T; 0 other;

Query Match 77.2%; Score 347.4; DB 22; Length 747;
 Best Local Similarity 97.0%; Pred. No. 3e-83;
 Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 atggatcgaagcagccagcagcagtcggtgttgcgttctccccaagaatcactgct 60
 |||
 Db 73 atgagatcgaagcagccagcagtcggtgttgcgttctccccaagaatcactgct 132
 OY 61 tcagaataatgagtcacagaacttcacgacagcacttcacccaagaagcccttgcaa 120
 |||
 Db 133 tcagaataatgagtcacagaacttcacgacagcacttcacccaagaagcccttgcaa 192
 OY 121 aaatatttgcagaagaatgaagaatcttaggactatccagactcctgttggaattatg 180
 |||
 Db 193 aaatatttgcagaagaatgaagaatcttaggactatccagactcctgttggaattatg 252
 OY 181 accctttcttggaggtatctcttcccttccctgtttaaaccatccaggtttccc 240
 |||
 Db 253 accctttcttggaggtatctcttcccttccctgtttaaaccatccaggtttccc 312
 OY 241 ttatattcttccagagatccatctcgtggctcgtgttgcgttcaatcctgagacc 300
 |||
 Db 313 ttatattcttccagagatccatctcgtggctcgtgttgcgttcaatcctgagacc 372
 OY 301 ttcccaattgcagtgaaaagaagaacacagaactctgggaatttgcattcatgtatg 360
 |||
 Db 373 ttcccaattgcagtgaaaagaagaacacagaactctgcataatttgagccgaataatg 432
 OY 361 acttt 365
 |||
 Db 433 aattt 437

RESULT 5
 ID AAS04281 standard; cDNA; 417 BP.
 AC AAS04281;

26-SEP-2001 (first entry)
 Novel human membrane protein #3 cDNA sequence.

Human; membrane protein; membrane receptor; IGF receptor; CD20;
 physiological disorder; ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 1..417
 FT /*tag= a
 FT /product= "Human membrane protein #3"

W0200146417-A2.

28-JUN-2001.

12-DEC-2000; 2000WO-US33742.

22-DEC-1999; 99US-0171567.

(LEXI-) LEXICON GENETICS INC.

Walke DW, Turner CA;

DR WPI: 2001-408646/43.
 DR P-PSDB: AA001212.
 XX
 PT Polynucleotide encoding novel human membrane protein, useful for
 PT identifying agonist, antagonist or modifiers or for producing
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic
 PT applications

Disclosure: Page 30; 32pp; English.

CC The present sequence encodes for novel human membrane protein #3.
 CC Human membrane protein #3 is 1 of 4 human membrane proteins
 CC (AA001210-AA001213) given in the present invention. These membrane
 CC proteins share structural similarity with membrane receptors such as
 CC the IGF receptor and mammalian CD20. The novel human membrane proteins
 CC are useful for identifying agonists, antagonists and modulators of the
 CC membrane proteins, and for producing antibodies specific to the
 CC membrane proteins. The membrane proteins can be used for diagnosis,
 CC drug screening, pharmacogenomic applications, clinical trial monitoring
 CC and the treatment of physiological disorders and diseases. The
 CC polynucleotides encoding the membrane proteins can be used to generate
 CC PCR primers or probes to identify mutations associated with a particular
 CC disease.

SQ Sequence 417 BP; 115 A; 91 C; 69 G; 142 T; 0 other;

Query Match 75.7%; Score 340.8; DB 22; Length 417;
 Best Local Similarity 94.1%; Pred. No. 1.5e-81;
 Matches 354; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 1 atggatcgaagcagccagcagtcggtgttgcgttctccccaagaatcactgct 60
 |||
 Db 1 atggatcgaagcagccagcagtcggtgttgcgttctccccaagaatcactgct 60
 OY 61 tcagaataatgagtcacagaacttcacgacagcacttcacccaagaagcccttgcaa 120
 |||
 Db 61 tcagaataatgagtcacagaacttcacgacagcacttcacccaagaagcccttgcaa 120
 OY 121 aaatatttgcagaagaatgaagaatcttaggactatccagactcctgttggaattatg 180
 |||
 Db 121 aaatatttgcagaagaatgaagaatcttaggactatccagactcctgttggaattatg 180
 OY 181 accctttcttggaggtatctcttcccttccctgtttaaaccatccaggtttccc 240
 |||
 Db 181 accctttcttggaggtatctcttcccttccctgtttaaaccatccaggtttccc 240
 OY 241 ttatattcttccagagatccatctcgtggctcgtgttgcgttcaatcctgagacc 300
 |||
 Db 241 ttatattcttccagagatccatctcgtggctcgtgttgcgttcaatcctgagacc 300
 OY 301 ttcccaattgcagtgaaaagaagaacacagaactctgggaatttgcattcatgtatg 360
 |||
 Db 301 ttcccaattgcagtgaaaagaagaacacagaactctgcataatttgagccgaataatg 360
 OY 361 acttcagcattatg 376
 |||
 Db 361 attccacaaaatag 376

RESULT 6
 ID AAC03080 standard; cDNA; 372 BP.
 AC AAC03080;

06-OCT-2000 (first entry)

Human secreted protein 5' EST, SEQ ID NO: 3078.

Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 gene therapy; chromosome mapping; ss.

OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GENSET).
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-500381/45.
XX
DR P-PSDB; AAG03074.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1: SEQ ID 3078; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
CC cDNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
XX

SQ Sequence 372 BP; 105 A; 87 C; 53 G; 117 T; 10 other;

Query Match 62.4%; Score 280.8; DB 21; Length 372;
Best Local Similarity 96.3%; Pred. No. 1.3e-65;
Matches 289; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 atgagattcaagaccgacacagtcggtgttcttggtattccctccgaataactgtc 60
|||||
DB 73 atgattcaagaccgacacagtcggtgttcttggtattccctccgaataactgtc 132
61 tcagatatgagtcacagaaacttcacgacagcctttcaactcaagcccttgca 119
|||||
DB 133 tcagatatgagtcacagaaacttcacgacagcctttcaactcaagcccttgca 192
QY 120 aaatattgtctagaaaataaactcttaggaactccagatccctgttgaattat 179
DB 193 aaatattgtctagaaaataaactcttaggaactccagatccctgttgaattat 252
QY 180 gaccttctcttggaagtattctcttcccttgtaaacatcatatccagtttcc 239
|||||
DB 253 gaccttctcttggaagtattctcttcccttgtaaacatcatatccagtttcc 312
QY 240 cttatattcttcagagatcattctgggctctgttctgtcattatctcggagc 299
|||||
DB 313 cttatattcttcagagatcattctgggctctgttctgtcattatctcggagc 372

RESULT 7
AAH64745
ID AAH64745 standard; cDNA; 468 BP.
XX
AC AAH64745;
XX
DT 11-SEP-2001 (first entry)

XX Human secreted protein cDNA, SEQ ID NO: 21.
DE
XX
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KM GENSET; ss.
XX
XX Homo sapiens.
OS
XX
PN WO200142451-A2.
XX
PD 14-JUN-2001.
XX
PF 07-DEC-2000; 2000MO-1B01938.
XX
PR 08-DEC-1999; 99US-0169629.
XX
PR 06-MAR-2000; 2000US-0187470.
XX
PA (GENSET).
XX
PI Dumas Milne Edwards J, Bougueleret L, Jobert S;
XX
DR WPI: 2001-367870/38.
XX
DR P-PSDB; AAG89142.
XX
PT Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases -
XX
XX Claim 7: Page 581; 921pp; English.

XX The invention relates to full length GENSET human nucleic acids encoding
XX potentially secreted proteins. The nucleic acids and the polypeptides
XX they encode may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate GENSET gene expression. For
XX example, they be used to treat disorders associated with decreased
XX GENSET gene expression by rectifying mutations or deletions in a
XX patient's genome that affect the activity of GENSET or by supplementing
XX the patient's own production of GENSET polypeptides. Conversely,
XX antisense nucleic acid molecules may be administered to down regulate
XX GENSET expression by binding with the cells' own genes and preventing
XX their expression. The sense and antisense nucleic acids may also be
XX used as DNA probes in diagnostic assays to detect and quantitate the
XX presence of similar nucleic acid sequences in samples, and hence to
XX determine which patients may be in need of restorative therapy.
XX The GENSET polypeptides may be used as antigens in the production of
XX antibodies and in assays to identify modulators (agonists and
XX antagonists) of GENSET polypeptide expression and activity. The
XX present sequence is a GENSET nucleic acid of the invention.

SQ Sequence 468 BP; 155 A; 94 C; 81 G; 138 T; 0 other;

Query Match 52.0%; Score 234; DB 22; Length 468;
Best Local Similarity 77.1%; Pred. No. 4.1e-53;
Matches 347; Conservative 0; Mismatches 0; Indels 103; Gaps 1;

QY 1 atgagattcaagaccgacacagtcggtgttcttggtattccctccgaataactgtc 60
|||||
DB 76 atgagattcaagaccgacacagtcggtgttcttggtattccctccgaataactgtc 135
QY 61 tcagatatgagtcacagaaacttcacgacagcctttcaactcaagcccttgca 120
|||||
DB 136 tcagatatgagtcacagaaacttcacgacagcctttcaactcaagcccttgca 195
QY 121 aaatattgtctagaaaataaactcttaggaactccagatccctgttgaattatg 180
|||||
DB 196 aaatattgtctagaaaataaactcttagg----- 227
QY 181 acccttctcttggaagtattctcttcccttgtaaacatcatatccagtttcc 240
DB 228 ----- 227
QY 241 ttatattcttcagagatcattctgggctctgttctgtcattatctcggagcc 300

```

Db 228 -----ggatatacattctggggtctgttctgttcttcaatttctgagcc 272
QY 301 ttccataatgcagtgtaaaagaaacacacagaacatctgggaatttgattacattgacg 360.
Db 273 ttccataatgcagtgtaaaagaaacacacagaacatctgggaatttgattacattgacg 332
QY 361 acttcaagattatgaattatcattctctgccttcttcaatttgggggagccactca 420
Db 333 acttcaagattatgaattatcattctctgccttcttcaatttgggggagccactca 392
QY 421 gaggattgtgattgtgaacaatgtgttga 450
Db 393 gaggattgtgattgtgaacaatgtgttga 422

```

RESULT 8

AA504280
ID AA504280 standard; cDNA; 246 BP.

AA504280;

26-SEP-2001 (first entry)

Novel human membrane protein #2 cDNA sequence.

Human; membrane protein; membrane receptor; IGF receptor; CD20; physiological disorder; ss.

Homo sapiens.

Key Location/Qualifiers
1..246
CDS /*tag= a
FT /product= "Human membrane protein #2"

WO200146417-A2.

28-JUN-2001.

12-DEC-2000; 2000WO-US33742.

22-DEC-1999; 99US-0171567.

(LEXI-) LEXICON GENETICS INC.

Walke DW, Turner CA;

WPI; 2001-408646/43.

P-PSDB; AAU01211.

Polynucleotide encoding novel human membrane protein, useful for identifying agonist, antagonist or modifiers or for producing antibodies useful in therapeutic, diagnostic and pharmacogenomic applications.

Disclosure; Page 30; 32pp; English.

The present sequence encodes for novel human membrane protein #2.

Human membrane protein #2 is 1 of 4 human membrane proteins (AAU01210-AAU01213) given in the present invention. These membrane proteins share structural similarity with membrane receptors such as the IGF receptor and mammalian CD20. The novel human membrane proteins are useful for identifying agonists, antagonists and modulators of the membrane proteins, and for producing antibodies specific to the drug screening, pharmacogenomic applications, clinical trial monitoring and the treatment of physiological disorders and diseases. The polynucleotides encoding the membrane proteins can be used to generate PCR primers or probes to identify mutations associated with a particular disease.

Sequence 246 BP; 60 A; 45 C; 47 G; 94 T; 0 other;

Query Match 25.1%; Score 113; DB 22; Length 246;
Best Local Similarity 100.0%; Pred. No. 5,6e-21;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 338 tgggaatttgattacattgagacttcaagcaattatgaattatcattctctgacct 397
Db 134 tgggaatttgattacattgagacttcaagcaattatgaattatcattctctgacct 193
QY 398 tctcaatttgggggtgacacacagagattgtgattgtgaacaatgtgttga 450
Db 194 tctcaatttgggggtgacacacagagattgtgattgtgaacaatgtgttga 246

```

RESULT 9

AA236238
ID AA236238 standard; cDNA; 1563 BP.

AA236238;

22-FEB-2000 (first entry)

cDNA encoding a bone marrow secreted protein designated BMS208.

Bone marrow secreted protein; bone marrow stromal cell; cytokine; cell proliferation; cell differentiation; hematopoiesis; anaemia; myeloid cell deficiency; lymphoid cell deficiency; myeloid cell; erythroid progenitor cell; colony stimulating factor; granulocyte; monocyte; macrophage; myelo-suppression; megakaryocyte; platelet; platelet disorder; thrombocytopenia; hematopoietic stem cell; stem cell disorder; aplastic anaemia; bone differentiation; paroxysmal nocturnal hemoglobinuria; bone growth; cartilage; tendon; ligament; nerve; wound healing; tissue repair; burn; incision; ulcer; bone fracture; cartilage damage; artificial joint; ss.

Homo sapiens.

Key Location/Qualifiers
255..758
CDS /*tag= a
FT /product= "bone marrow secreted protein"

sig_peptide 255..440

polya_signal 1531..1536

/*tag= c

WO9933979-A2.

08-JUL-1999.

18-DEC-1998; 98WO-US27008.

30-DEC-1997; 97US-0068958.

24-SEP-1998; 98US-0101603.

30-SEP-1998; 98US-0102540.

(CHIR) CHIRON CORP.

Liu H, Cao L;

WPI; 2000-038344/03.

P-PSDB; AAY53632.

New isolated human polynucleotide and secreted proteins can induce production of other cytokines in certain cell populations.

Claim 11; Page 92-94; 120pp; English.
AA236228-49 encode bone marrow secreted proteins of human bone marrow stromal cells. The proteins can exhibit cytokine, cell proliferation, or cell differentiation activity (either inducing or inhibiting). They can be used to support colony forming cells or factor-dependent cell lines,

XX	24-APR-2001	(first entry)	
DF			
XX	Oligonucleotide D2004.		
DE			
XX	Electron-transfer group: ETM; mismatch; genotyping;		
KW	gene expression; ss.		
KW			
XX	Synthetic.		
OS			
XX	WO200107665-A2.		
PN			
XX	01-FEB-2001.		
XX			
FE	26-JUL-2000; 2000WO-US20476.		
PR	26-JUL-1999; 99US-0145695.		
PR	17-MAR-2000; 2000US-0190259.		
XX			
PA	(CLIN-) CLINICAL MICRO SENSORS INC.		
XX			
XX	Umek RM;		
DR	WPI; 2001-159728/16.		
XX			
PT	Nucleic acids containing electron-transfer group, useful as labels in		
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on		
PT	a single surface		
XX			
PS			
XX	Example 6; Page 128; 159pp; English.		
CC			
CC	The present invention relates to a composition comprising two nucleic		
CC	acids each containing an electron-transfer group (ETM) having		
CC	different redox potentials. The invention is used for electronic		
CC	detection of nucleic acids, especially of substitutions (mismatches)		
CC	and single-nucleotide polymorphisms, e.g. for genotyping,		
CC	monitoring gene expression.		
XX			
XX			
SO	Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;		
	Query Match	16.4%; Score 73.6; DB 22; Length 936;	
	Best Local Similarity	1.7%; Pred. No. 2.6e-10;	
	Matches	7; Conservative 261; Mismatches 150; Indels 0; Gaps	0;
OY	31 ttctcgtattctcccgagaatcactcgtctcagaatatgagtcacagaacttcagcc	90	
DB	75 www.	134	
OY	91 acgaccttcaactcaagaagcccttgcaaaatattgtctagaaaatgaaatctta	150	
DB	135 www.	194	
OY	151 gggactacacagatccgttttggaattagaccttcttcttggaattatcccttcc	210	
DB	195 wgw.	254	
OY	211 acctcgttaaaacacatcacagaagttcccttataattcttccagagataccattc	270	
DB	235 www.	314	
OY	271 ggcctcgtttgtcattaatctcgaagccttcctcaatgacgtgaaagaaaacaca	330	
DB	315 www.	374	
OY	331 gaaactcgggaatttgattacattgacgttcacagcatatgaattatcattct	390	
DB	375 www.	434	
OY	391 ctgccttccatatttgggggctccacaccagagattgagttggaacaatggtgt	448	
DB	435 www.	492	

```

RESULT 14
AAF58262
ID AAF58262 standard; DNA; 936 BP.
XX
AC AAF58262;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2007.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Unkek RM;
XX
WP1: 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
PS Example 6; Page 128; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;

Query Match      16.4%; Score 73.6; DB 22; Length 936;
Best Local Similarity 1.7%; Pred. No. 2,6e-10;
Matches 7; Conservative 261; Mismatches 150; Indels 0; Gaps 0;

QY 31 ttctcgatattcccccgaataactcgtctagaatatgagtcacagaaattccaagcc 90
   : :::: : : :::: : : : :::: : : : :::: : : : :::: :
DB 75 www.wwwww.wwwww.wwwww.cwww.wwwww.wwwww.wwwww.wwwww.wwwww 134
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 91 agaaccttcaactcaagcccttgcaaaaataattctgcagaanaagtgaatcta 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 135 www.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 151 gggactatcacatccgtttggaattagaccttctttggagttactctccttc 210
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 195 wgvwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 211 acctgttaaaccatocaaagtttcccttatattcttcaggataccattcgg 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 255 www.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww 314
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 271 ggctctgttttgtaatatctcggagccctcctaattgcaglygaanaagaaacca 330
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 315 www.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww.wwwww 374
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 331 gaactctgggaatttgattacatgatgacttcagcatatgatgaattatcattct 390
   : : : : : : : : : : : : : : : : : : : : : : : : : : :
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2001, 21:39:28 : Search time 83.78 Seconds
(without alignments)
1216.461 Million cell updates/sec

Title: US-09-735-712-7

Perfect score: 450
Sequence: 1 atgtattcaagcaccgcaca.....atgtgacaatgtgttga 450

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	ID	Description
No.						
1	66.8	14.8	1669	2	US-08-916-902A-2	Sequence 2, Appl
2	66.8	14.8	1669	2	US-09-213-389-2	Sequence 2, Appl
3	46	10.2	1661	1	US-08-318-492-3	Sequence 3, Appl
4	46	10.2	1661	1	US-08-707-340-3	Sequence 3, Appl
5	46	10.2	1661	2	US-08-994-578-3	Sequence 3, Appl
6	44.4	9.9	2545	1	US-07-869-933-22	Sequence 22, Appl
7	44.4	9.9	2545	4	US-09-103-663-22	Sequence 22, Appl
8	41.6	9.2	7218	1	US-08-232-463-14	Sequence 14, Appl
9	34.8	7.7	3393	4	US-09-104-324B-1	Sequence 1, Appl
10	34.8	7.7	3393	4	US-09-162-713-1	Sequence 1, Appl
11	33.6	7.5	1711	1	US-08-568-147B-1	Sequence 1, Appl
12	33.4	7.4	1185	2	US-08-179-557-14	Sequence 14, Appl
13	33.4	7.4	6669	3	US-09-212-971-5	Sequence 5, Appl
14	33.4	7.4	6669	3	US-08-800-929A-5	Sequence 5, Appl
15	33.4	7.4	6669	3	US-09-617-053A-5	Sequence 5, Appl
16	32.6	7.2	636	2	US-08-284-941-3	Sequence 3, Appl
17	32.6	7.2	636	2	US-08-447-642-3	Sequence 3, Appl
18	32.6	7.2	636	4	US-09-236-503-3	Sequence 3, Appl
19	32.6	7.2	636	5	PCR-US93-02147A-3	Sequence 3, Appl
20	31.6	7.0	5816	4	US-09-220-641-4	Sequence 4, Appl
21	31.4	6.9	4517	4	US-09-140-804-9	Sequence 9, Appl
22	31	6.9	2839	4	US-09-061-702-1	Sequence 1, Appl
23	30.8	6.8	6699	5	PCR-US95-13335-2	Sequence 2, Appl
24	30.8	6.8	2662	3	US-08-750-357-8	Sequence 8, Appl
25	30.4	6.8	591	4	US-09-385-983-300	Sequence 300, App
26	30.4	6.8	2189	3	US-08-846-020A-1	Sequence 1, Appl
27	30.4	6.8	4699	1	US-08-181-271A-6	Sequence 6, Appl

28	30.4	6.8	4699	1	US-08-449-315-6	Sequence 6, Appl
29	30.4	6.8	4699	1	US-08-444-803-6	Sequence 6, Appl
30	30.4	6.8	4699	1	US-08-449-043-6	Sequence 6, Appl
31	30.4	6.8	4699	1	US-08-456-265A-6	Sequence 6, Appl
32	30.4	6.8	4699	1	US-08-455-416-6	Sequence 6, Appl
33	30.4	6.8	4699	1	US-08-455-244-6	Sequence 6, Appl
34	30.4	6.8	4699	1	US-08-454-876-6	Sequence 6, Appl
35	30.4	6.8	4699	2	US-08-457-364-6	Sequence 6, Appl
36	30.4	6.8	4699	2	US-08-456-262-6	Sequence 6, Appl
37	30.4	6.8	4699	2	US-08-456-240-6	Sequence 6, Appl
38	30.4	6.8	4699	2	US-08-455-736-6	Sequence 6, Appl
39	30.4	6.8	4699	2	US-08-971-217-6	Sequence 6, Appl
40	30.4	6.8	4699	4	US-09-350-600-6	Sequence 6, Appl
41	30.4	6.8	22846	2	US-08-469-461-3	Sequence 3, Appl
42	30.4	6.8	22846	3	US-07-890-609-3	Sequence 3, Appl
43	30.2	6.7	315	2	US-08-743-200-1	Sequence 1, Appl
44	30.2	6.7	376	2	US-08-623-906A-18	Sequence 18, Appl
45	30.2	6.7	2529	2	US-08-883-795A-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-08-916-902A-2
Sequence 2, Application US/08916902A
Patent No. 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,902A
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINF04
CLONE: 927955
US-08-916-902A-2

	Query Match	14.8%	Score 66.8	DB 2	Length 1669	
	Best Local Similarity	53.4%	Pred. No. 1,3e+09			
	Matches	140	Conservative	0	Mismatches 122	Indels 0
					Gaps	0
QY	119	aaaaattattgctagaanaaaatcttaggagactatccagatccctgtttgaatla	178			
Db	325	AGAACTCTTTGAAAGGAGAGAACCCAAAGTCCTTGGGGGTTGTGCAGATTCCTGACCTGCCCTGA	384			
QY	179	tgaaccttctctttagagatcatctcccttacccttgcctgtttaaaccataaccagaatttc	238			
Db	385	TGAGCCTTAGCATGGGAAATTAACAATGATGTGTATGGCATCTAATTACTTATGGAAGTAAC	444			
QY	239	cccttaattcttcttcaagatatccatctcgtggcgtctgtttgttcatlaattcctgag	298			
Db	445	CTATTTCGGTATATTCGGGTACAACAATTTGGGGGTCATATATCTTATTAATTTACGAT	504			
QY	299	ccctccaaattgcagtgaanaagaaaacacagaaactcttggaatttggattacatga	358			
Db	505	CCTTGTCAAATTGCACAGCAAGAAATTAGACACTACAAAAGGCGCTGTCGAGGATATCTTAGGAA	564			
QY	359	tgaccttcaagcatatttgaatt	380			
Db	565	TGAATATCACCAGCTCTGTACT	586			

RESULT 2
 US-09-213-389-2
 : Sequence 2, Application US/09213389
 : Patent No. 5970702
 :
 : GENERAL INFORMATION:
 : APPLICANT: Bandman, Olga
 : APPLICANT: Lal, Preeti
 : APPLICANT: Corley, Neil C.
 : TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
 : TITLE OF INVENTION: RECEPTOR-LINE PROTEIN
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Incyte Pharmaceuticals, Inc.
 : STREET: 3174 Porter Drive
 : CITY: Palo Alto
 : STATE: CA
 :
 : COUNTRY: USA
 : ZIP: 94304
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS
 :
 : SOFTWARE: FASTSEQ for Windows Version 2.0
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/213,389
 : FILING DATE:
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/916,902
 : FILING DATE:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Billings, Lucy J.
 : REGISTRATION NUMBER: 36,749
 : REFERENCE/DOCKET NUMBER: PE-0371 US
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 415-855-0555
 : TELEFAX: 415-845-4156
 : TELEX:
 :
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1669 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : IMMEDIATE SOURCE:
 : LIBRARY: BRAINOT04
 : CLONE: 927955
 :
 : US-09-213-389-2

Query Match	14.88;	Score 66.8;	DB 2;	Length 1669;
Best Local Similarity	53.44;	Pred. No. 1.3e-09;		
Matches 140;	Conservative 0;	Mismatches 122;	Indels 0;	Gaps
Qy 119	aaaatttttgcgcagaaaaatgaaatccttagggactatccagatccctgttggaa	178		
Db 325	AGAACTTTTGAAAGGAGACCCAAAGCTTGGGGTTGTGCACAAATTCGACTCCCTGA	388		
Qy 179	tgaccttctcttggaggttaacctcttccattacacgtttaaaccalatcaaggttcc	238		
Db 385	TGAGCCTTAGAGATGGGAATTAACAATGATGTGTATGGCATCTAAATCTTAGGAAGTAACC	444		
Qy 239	ccctataattctcttcagagatccatccatctcgggctcgtgttcttcaattaattcggag	298		
Db 445	CTATTCCGCTATATATCGGGTACCAATTTGGGGTGATGATGTTTATTTTACAGAT	504		
Qy 299	ccctcccaattgcaggtgtaaaaaaaccacaagaactctcgggaatttgattacatga	354		
Db 505	CCCTGTCAATTGACAGCAGCAATTTGAACTACAAAAAGCGCTGGTCCGAGGTAGTCTAGGAA	564		
Qy 359	tgaccttcagcatattgaatt	380		
Db 565	TGAATATCCACGCTCTTACT	586		

```

1      RESULT 3
2      US-08-318-492-3
3      Sequence 3, Application US/08318492
4      Patent No. 5552312
5      GENERAL INFORMATION:
6      APPLICANT: Lim, Bing
7      APPLICANT: Adra, Chaker N.
8      APPLICANT: Lellias, Jean-Michel
9      TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
10     TITLE OF INVENTION: ASSAYS
11     NUMBER OF SEQUENCES: 4
12     CORRESPONDENCE ADDRESS:
13     ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
14     STREET: Two Miltitia Drive
15     CITY: Lexington
16     STATE: Massachusetts
17     COUNTRY: U.S.A.
18     ZIP: 02173
19     COMPUTER READABLE FORM:
20     MEDIUM TYPE: Floppy disk
21     COMPUTER: IBM PC compatible
22     OPERATING SYSTEM: PC-DOS/MS-DOS
23     SOFTWARE: PatentIn Release #1.0, Version #1.25
24     CURRENT APPLICATION DATA:
25     APPLICATION NUMBER: US/08/318,492
26     FILING DATE:
27     CLASSIFICATION: 530
28     ATTORNEY/AGENT INFORMATION:
29     NAME: Brook, David E.
30     REGISTRATION NUMBER: 22,592
31     REFERENCE/DOCKET NUMBER: B1H94-03
32     TELECOMMUNICATION INFORMATION:
33     TELEPHONE: (617) 861-6240
34     TELEFAX: (617) 861-9540
35     INFORMATION FOR SEQ ID NO: 3:
36     SEQUENCE CHARACTERISTICS:
37     LENGTH: 1661 base pairs
38     TYPE: nucleic acid
39     STRANDEDNESS: single
40     TOPOLOGY: linear
41     MOLECULE TYPE: DNA (genomic)
42     FEATURE:
43     NAME/KEY: CDS
44     LOCATION: 97..741
45     US-08-318-492-3

```

Query Match	10.2%	Score 46;	DB 1;	Length 1661;
Best Local Similarity	55.2%;	Pred. No. 0.00074;		
Matches 116; Conservative	0;	Mismatches 85;	Indels 9;	Gaps 1

QY	135	aaaaatgaaatccttaggactaccagatccctgttggaaatcttgaccttctttg	134
Db	234	AAATTTACAACTTTCTTGGGGGCATCCAGATCCCTAATGACACAAATGATTTGGCTTTGGG	233
QY	195	agt-----tatcttcctttacccctgtgtaaaacacatccaaagtttcccttat	245
Db	294	TGTTCTTCTGGGCTTCGCAATTCACATTCACAACTTCCAAAAGACACTTTCTTTCTTCAC	353
QY	246	attcctcttcaggatccatctctgggagctctgttttcttcaataatctcggagcttccct	305
Db	354	CTTTACACAGGCTACCCCGATTTGGGGTGCCTGTCTTTTCTGTATGTTACAGAAACCTTGTC	413
QY	306	aattgcagtgaaaagaaaacacaganaac	335
Db	414	TGTTGTACAGGGATTAATAACCCACAAAGAC	443

RESULT 4

US-08-707-340-3
; Sequence 3, Application US/08707340
; Patent No. 5705615
; GENERAL INFORMATION:

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994

ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03A2

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

STRANDEDNESS: single

MOLECULE TYPE: DNA (genomic)

```

; NAME/KEY: CDS
; LOCATION: 97..738

```

US-08-707-340-3

Query Match	10.2%;	Score 46;	DB 1;	Length 1661;
Best Local Similarity	55.2%;	Pred. No. 0.00074;		
Matches 116;	Conservative 0;	Mismatches 85;	Indels 9;	Gaps 1.

QY	135	aaaatgaaacctcttaggacatccagatccctcttggaaatctgaccttcctttg	194
Db	234	AAAATTACAACTTTCTTGGGGGCATCCAGATCCCTGAATGACGACAAATGATTCCTGGCTTTGGG	293
QY	195	agt-----tatcttccttttacctctgttaaaacatccatccaaagttcccttat	245
Db	294	TGTCCTTCTGGGCTCCCTGCACAAATACCATTAACACTTCACAAAGACCTTTCTTCTTCAC	353
QY	246	attctcttcaggaaatccatctcggggcctctgttttcttaataatctcggagccttcc	305
Db	354	CTTCTACACAGGCTCAACCCGATTTGGGTGGCTGTCTTTTCTGTATGTCAGGAACCTTGTG	413
QY	306	aattgcagtgaaagaaanaaacacagaaac	335
Db	414	TGTTGTAGCAGGGATTAATAACCCACAAAGAC	443

RESULT 5

US-08-994-578-3
; Sequence 3, Application US/08994578
; Patent No. 5972688

: APPLICANT: Lim, Bing
 : APPLICANT: Adra, Chaker N.
 : APPLICANT: Leijas, Jean-michel
 : TITLE OF INVENTION: HYDRA METHODS OF TREATMENT AND ASSAYS
 : TITLE OF INVENTION: AGONISTS AND ANTAGONISTS

```

:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: PatentIn Release #1.0, Version #1.300
:

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/707,340
FILING DATE: 03-SEP-1996

APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
ATTORNEY/AGENT INFORMATION:

TELECOMMUNICATION INFORMATION
TELEPHONE: (781) 861-6240

SEQUENCE CHARACTERISTICS:

TYPE: nucleic acid

FEATURE:

Db 308 cttcagacttgcagcagaagtgctttatattatagagcagcgtaccattctgggtg 367
Qy 275 ctgttttcttcaatccttgagccctcctaattgcagtgaaagaacacacagaa 334
Db 368 cagtgctgttctgtctgtgattttgtcaattatgtccgaagaagaacacactgt 427
Qy 335 cctcgtg 340
Db 428 atctcgtg 433

RESULT 8
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESS: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F15
; US-08-232-463-14

Query Match 9.2%, Score 41.6; DB 1; Length 7218;
Best Local Similarity 4.2%; Pred. No. 0.02; Mismatches 173; Indels 0; Gaps 0;
Matches 17; Conservative 214; Mismatches 173; Indels 0; Gaps 0;

Qy 6 ttcaagcaccgcacacagtcgcgtgttctgttattctcctcagaatacactgcttcaga 65
Db 1032 TTCGAGCTTGCGTGCAGTGCAGGAGCTGTGCGATYYYYYYYYYYYYYYYYYYY 1091
Qy 66 atatagtcacagaacttgcagccagcactttcaactcaaaagccctcgtcaaaatt 125
Db 1092 YY 1151

Qy 126 atttgcagaataatccttaggcactatccagatccctgtttgaaattgacctt 185
Db 1152 YY 1211
Qy 186 ttctttgagtaatccttcttcaacttgtaaacacataccaagttcccttat 245
Db 1212 YY 1271
Qy 246 attcttcaagatatccactctgggctcgttttcttcaatcaattctgagccctcct 305
Db 1272 YY 1331
Qy 306 aattgcagtgaaagaataaacacagaaactctgggaatttgattcaattgacttt 365
Db 1332 YY 1391
Qy 366 cagcatattgaattatcattcctcctccttcccaatttgg 409
Db 1392 YY 1435

RESULT 9
US-09-104-324B-1/c
; Sequence 1, Application US/09104324B
; Patent No. 6232460
; GENERAL INFORMATION:
; APPLICANT: T reci, Ozlem, Sahin, Ugur, Pfrendschuh, Michael
; TITLE OF INVENTION: Methods For Diagnosis And Treating Cancers,
; TITLE OF INVENTION: And Methods For Identifying Pathogenic Markers-In A Sample
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fulbright & Jaworski LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104.324B
; FILING DATE: 25-June-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/892,702
; FILING DATE: 15-July-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6232460man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LOD 5491
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-104-324B-1

Query Match 7.7%, Score 34.8; DB 4; Length 3393;
Best Local Similarity 49.5%; Pred. No. 1.2;
Matches 90; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 186 ttctttgagtaatccttcttcaacttgtaaacacataccaagttcccttat 245
Db 3004 TTCCTTAGATTCTTTTTCATTCATTATGCAATTAACAGCCCAACGCTCCTCCGCA 2945

OY	246	atctctctcaggaatcattctcgggacgtcgtttgtgtcatcaatctcgagccctctc	305
Db	2944	TTTTCTATAGCTCCAAACTTCAGTGTAGTCCAGGCGTTGTTCAGATGAAAGGCGCTT	28855
OY	306	aattgcgcgtgaaaaagaaaacccagaacctcgtgggaatttgatcacatgatgaactt	365
Db	2884	TTTTCGTCGTCTTTCACACAAAGATGACGAAGCTGTGTGATTATGTCCTATACAGTGTCTT	28235
OY	366	ca	367
Db	2824	CA	2823

RESULT 10
 US-09-162-713-1/c
 Sequence 1, Application US/09162713
 Patent No. 6239256
 GENERAL INFORMATION:
 APPLICANT: T reel, Ozlem; Sahin, Ugur; Pfrendschuh, Michael
 TITLE OF INVENTION: Methods For Diagnosis And Treating
 TITLE OF INVENTION: Cancers, And Methods For Identifying Pathogenic Markers In A
 SAMPLE OF INVENTION: Sample Of No. 6239256mml Cells
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Felfe & Lynch
 STREET: 805 Third Avenue
 CITY: New York City
 STATE: New York
 COUNTRY: USA
 ZIP: 10022
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
 COMPUTER: IBM
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: Wordperfect
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/162.713
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/892,702
 FILING DATE: 15-July-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Hanson, No. 6239256man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5491
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3393 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-09-162-713-1

	Query Match	7.7%;	Score 34.8;	DB 4;	Length 3393;
	Best Local Similarity	49.5%;	Pred. No. 1.2;		
	Matches 90;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0;
Oy	186 ttctcttgaggtatcttcctttcaccttgyttaaacatatccaagtcccttat	245 			
Dd	3004 TTTCTTATAGTTTTTTTTTCATCCATTATGCAATTACAGGCCAACGGCTCCTCCCGCAT	2945 			
Oy	246 attcttcacagatatacatcctcgsggctcgtttgtlcaataattctgtagccttctc	305 			
Dd	2944 TTTTCTTATAGCTCCAACACTTGACGTAGGTGCAGGGGTTGTAAAGATGAAGGGGCTT	2885 			
Oy	306 aattgcagtgaaagaanaaacaccagaacctcgsgaatlttgtatcaatgatgacttt	365 			

Db	2884	TTTTGGTGTTCACACCAAGAATGAGAACTGTGATTTATTTGTTCTATACAGCTTT	2825
QY	366	ca 367	
	2824	CA 2823	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	7.58;	Score 33.6;	DB 1;	Length 1711;	
	Best Local Similarity	47.68;	Pred. No. 2;			
	Matches	99;	Conservative	0;	Mismatches 109;	Indels 0; Gaps 0;
OY	101 caactcaaaagcccttcgcaaaaattatttgcgtagaanaatgyaaaactcttgaggactatcc	160				
Db	425 CATCTAATATTTCCTGTAATAATAGATACGTTTTTCGCCAGAAGCGTTGTTGGCGACTC	366				
OY	161 agaccgcgtttggagaataatgaacctttccttttggagtatcttcctttcaccttgttaa	220				
Db	365 TCACGCTCATATGCATTCAGTGCGATTGTGTGAATTCAGTCAGAAGCGTTTGAACAACGTGTCAT	306				
OY	221aacatatccaaggttcccttattatcttcctaagaatarccatctctggcgctctgttt	280				

Db	Qy
305	281
GACATTTTCCGACCCATCAACAAGATAGTTGGAGCCTTTCCGTGTGTCTCTGTGA	tgctcataatctcggagcctctcta
245	218
CTTGATCAAAAGTGAAGAACCTTTCTTAAT	cttgatcaaaagtgaagaaacctttctta

RESULT 12
 US-08-179-557-14/c
 : Sequence 14, Application US/08179557
 : Patent No. 5837509
 : GENERAL INFORMATION:
 : APPLICANT: ISRAELSEN, Hans
 : APPLICANT: BECH HANSEN, Egon
 : APPLICANT: MADSEN, Soeren Michael
 : APPLICANT: JOHANSEN, Eric
 : APPLICANT: NILSSON, Dan
 : APPLICANT: VRANG, Astrid
 : TITLE OF INVENTION: Recombinant Lactic Acid Bacterium
 : TITLE OF INVENTION: Containing an Inserted Promoter and Method of Constructing
 : TITLE OF INVENTION: Same
 : NUMBER OF SEQUENCES: 39
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: Foley & Lardner
 : STREET: 3000 K Street, N.W.
 : CITY: Washington, D.C.
 : COUNTRY: USA
 : ZIP: 20007-5109
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/179,557
 : FILING DATE: 07-JAN-1994
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: DK 1579/92
 : FILING DATE: 30-DEC-1992
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: DK 0988/93
 : FILING DATE: 01-SEP-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/DK94/00004
 : FILING DATE: 03-JAN-1994
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/036,681
 : FILING DATE: 25-MAR-1993
 : ATTORNEY/AGENT INFORMATION:
 : NAME: BENT, Stephen A.
 : REGISTRATION NUMBER: 29,768
 : REFERENCE/DOCKET NUMBER: 30307/140/PLVI
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 202 672 5300
 : TELEFAX: 202 672 5399
 : TELEX: 904136
 : INFORMATION FOR SEQ ID NO: 14:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1185 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : US-08-179-557-14

Query March	7.48:	Score 33.4:	DB 2,	Length 1185:
Best Local Similarity	54.5%:	Pred. No. 2:		
Matches 67,	Conservative	0;	Mismatches 56;	Indels 0;
			Gaps	0
OY	284	tcataatattctgagccttcctaattgcgcygaaagaaaaacacagaaactctggaa	343	-
db	421	TAATGAGTTTAAACCGCATCATCATGGAATTTAAATTCAGAGTCAGTCAGTACATCGCAG	362	

QY	344	tttggatcaacttgatgcgttcgcgaattatgaaattcaatcctgcgtttccaa	403
Dd	361	CAATATACAGATTCTGTGACTTTTCCTTTTGTGTAGAGCATTTGAGTTTCCCTTTCTT	302
QY	404	ttt 406	
Dd	301	TAT 299	

```

RESULT 13
US-09-212-971-5/c
; Sequence 5, Application US/09212971B
; Patent No. 6107041
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPs AND
; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; FILE REFERENCE: 07891/009002
; CURRENT APPLICATION NUMBER: US/09/212,971B
; CURRENT FILING DATE: 1998-12-16
; EARLIER APPLICATION NUMBER: 60/017,354
; EARLIER FILING DATE: 1996-04-26
; EARLIER APPLICATION NUMBER: 60/030,590
; EARLIER FILING DATE: 1996-11-14
; EARLIER APPLICATION NUMBER: 08/800,929
; EARLIER FILING DATE: 1997-02-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 6669
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variation
LOCATION: (3677)..(3951)
OTHER INFORMATION: n can be any nucleotide
US-09-212-971-5

```

[illegible]

RESULT 14
 US-08-800-929A-5/c
 Sequence 5, Application US/08800929A
 Patent No. 6133437
 GENERAL INFORMATION:
 APPLICANT: Kornelius, Robert G.
 APPLICANT: Mackenzik, Alexander
 APPLICANT: Liston, Peter
 APPLICANT: Baird, Stephen
 APPLICANT: Tsang, Benjamin K

```

; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF
; TITLE OF INVENTION: IAPS AND NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESS: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,929A
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,590
; FILING DATE: 14-NOV-1996
; APPLICATION NUMBER: 60/017,354
; FILING DATE: 26-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 07891/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-08-800-929A-5

```

```

Query Match          7.4%; Score 33.4; DB 3; Length 6669;
Best Local Similarity 53.4%; Pred. No. 3.6;
Matches 70; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

```

```

128 ttgctagaataatgaataatccttaggactatccagatcctgttggaaattatgaccttt 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2946 TGGCTGTAAATTAAGACCTTCAAGATTATTAATTTTGTGCCCCCTCAAGAACTAAC 2887
OY 188 ctttggagttatcttcccttccaccttgtaaacatccatccagaagtttcccttat 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2886 ATTGTAGAGTTATTCAACTTATCACCCTCTGAAACATTTTGAACCTTTGTTTTCRA 2827
OY 248 ttcttcagga 258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2826 ATGTTTCTGAA 2816

```

```

RESULT 15
US-09-617-053A-5/c
; Sequence 5, Application US/09617053A
; Patent No. 6300492
; GENERAL INFORMATION:
; APPLICANT: Korneluk, Robert G
; APPLICANT: Mackenzie, Alexander E
; APPLICANT: Liston, Peter
; APPLICANT: Baird, Stephen
; APPLICANT: Tsang, Benjamin K
; APPLICANT: Pratt, Christine
; TITLE OF INVENTION: DETECTION AND MODULATION OF IAPS AND

```

```

; TITLE OF INVENTION: NAIP FOR THE DIAGNOSIS AND TREATMENT OF PROLIFERATIVE
; TITLE OF INVENTION: DISEASE
; FILE REFERENCE: 07891/009003
; CURRENT APPLICATION NUMBER: US/09/617,053A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/800,929
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (3677)...(3951)
; OTHER INFORMATION: n can be any nucleotide
US-09-617-053A-5

```

```

Query Match          7.4%; Score 33.4; DB 4; Length 6669;
Best Local Similarity 53.4%; Pred. No. 3.6;
Matches 70; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

```

```

OY 128 ttgctagaataatgaataatccttaggactatccagatcctgttggaaattatgaccttt 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2946 TGGCTGTAAATTAAGACCTTCAAGATTATTAATTTTGTGCCCCCTCAAGAACTAAC 2887
OY 188 ctttggagttatcttcccttccaccttgtaaacatccatccagaagtttcccttat 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2886 ATTGTAGAGTTATTCAACTTATCACCCTCTGAAACATTTTGAACCTTTGTTTTCRA 2827
OY 248 ttcttcagga 258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2826 ATGTTTCTGAA 2816

```

```

Search completed: November 29, 2001, 00:20:23
Job time: 9655 sec

```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2001, 21:34:03 ; Search time 1970.99 Seconds
(without alignments)
2453.387 Million cell updates/sec

Title: US-09-735-712-7
Perfect score: 450
Sequence: 1 atgtatcaagcaccgcacac.....atgtgacaatgtgttga 450

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 segs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estrom:*
5: em_estrpl:*
6: em_estrda:*
7: em_estro:*
8: em_estrov:*
9: em_hic:*
10: qb_est1:*
11: qb_est2:*
12: qb_hic:*
13: qb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
C 1	370	82.2	410 10	AA411806 zt67a03.s
C 2	348	77.3	389 10	AA781801 a158901.s
C 3	347.4	77.2	516 10	AA436088 zu03a08.r
C 4	215.4	47.9	508 10	A1149899 qf43h06.x
C 5	214.8	47.7	482 10	A1002083 ol38b02.s
C 6	197.2	43.8	281 10	AA707529 ah41a12.s
C 7	157.8	33.1	415 10	AA470059 zt94h05.r
C 8	145.8	32.4	484 10	BE107659 UT-R-BT1-
C 9	126.2	28.0	224 10	BE638325 EST00022
C 10	113	25.1	382 10	AA435988 zu03a08.s
C 11	113	25.1	387 10	AA416972 zt94h05.s
C 12	112.4	25.0	538 13	AO108532 C17-HSP-2

Result No.	Score	Query Match Length	DB ID	Description
C 13	94.6	21.0	394 10	AA758635 ah67b04.s
C 14	83	18.4	218 10	BE638321 EST00013
C 15	70.6	15.7	421 13	B86842 RPT11-26F1
C 16	69.8	15.5	857 11	BE720182 602592226
C 17	68.4	13.2	534 11	BE591089 zt52902.x
C 18	66.8	14.8	508 10	AA234138 zt51b06.r
C 19	66.8	14.8	517 11	BE664654 DRABG07
C 20	66.8	14.8	704 11	BE571626 602592934
C 21	66.8	14.8	793 11	BE484817 602505511
C 22	66.8	14.8	1071 10	AL544561 AL544561
C 23	63.2	14.5	562 10	AA18443 zv92e05.r
C 24	58	12.9	494 11	BF475316 7i07h06.x
C 25	57.2	12.7	689 11	BE719295 602690369
C 26	57	12.7	899 11	BE720623 602691889
C 27	55.6	12.4	601 10	BE638317 EST00003
C 28	53.2	11.8	715 11	BE538851 602568296
C 29	52.4	11.6	344 11	BE861164 UT-M-AM0-
C 30	49	10.9	470 10	AL807884 w43911.x
C 31	48.2	10.7	1101 13	CNS0182P
C 32	47.8	10.6	1007 13	CNS06X9S
C 33	47.4	10.5	557 13	AO580877 RPT-11-4
C 34	47.2	10.5	332 11	BF319786 uv63h11.x
C 35	46.4	10.3	944 13	CNS06WVA
C 36	46	10.2	887 10	AL531049
C 37	45.6	10.1	1101 13	CNS000B8
C 38	44.8	10.0	1001 13	CNS01400
C 39	44.2	9.8	668 10	AL514901
C 40	44	9.8	1101 13	CNS0039G
C 41	43.8	9.7	647 13	B10659
C 42	43.4	9.6	888 13	CNS0170M
C 43	43.4	9.6	1101 13	CNS0039L
C 44	43.2	9.6	1101 13	CNS016HF
C 45	43	9.6	816 11	BE495708 602540318

ALIGNMENTS

RESULT 1
LOCUS AA411806 410 bp mRNA
DEFINITION zt67a03.s1 Soares_testis_NRT Homo sapiens cDNA clone IMAGE:777372
3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;
mRNA sequence.

ACCESSION AA411806
VERSION AA411806.1 GI:2070377

KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 410)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Getzel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.

TITLE WashU-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

FEATURES
source
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 608 Std Error: 0.00
Seq primer: -41m3 fwd. Et from Amersham
High quality sequence stop: 401.
Location/Qualifiers
1. 410

/organism="Homo sapiens"
/db_xref="GDB:5924284"

REFERENCE 1 (bases 1 to 516)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, T., Wylie, T., Waterston, R., and Wilson, R.
 TITLE WashU-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 733 Std Error: 0.00
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 482.

FEATURES
 source
 1..516
 /organism="Homo sapiens"
 /db_xref="GDB:5927437"
 /db_xref="taxon:9606"
 /clone="IMAGE:730742"
 /clone.lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCGAATGGAGGCGGCCCAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaído."

BASE COUNT 149 a 117 c 80 g 170 t
 ORIGIN

Query Match 77.2%; Score 347.4; DB 10; Length 516;
 Best Local Similarity 97.0%; Pred. No. 4.1e-68;
 Matches 354; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 atgagatcaagaacgcgacagtcgcgtgttctggtatttcctccagaatacactgct 60
 |||
 DB 52 ATGGATTCAAGACACCGACAGTCGGTGTGCTATTTCTCCAGAAATCACTGCT 111
 |||
 Y 61 tcagaatagatgcacagaacttcagcacgaactttcaactcaagcccttgcaa 120
 |||
 DB 112 TCAGAAATAGATCCACAGAACTTTCAGCCAGACCTTTTCAACTCAAGCCCTTGCA 171
 |||
 QY 121 aaatattgctagaaaaatgaatacttagagactatccagatccctgttggaaattg 180
 |||
 DB 172 AAATATTGTGTAGAAATGAATCTTAGGACTATCCAGATCCTGTTGGAATTATG 231
 |||
 QY 181 accttctcttggagatctctctcttcaacctgttaaacacatccaaagttcc 240
 |||
 DB 232 ACCTTTCTTTGGAGTATCTCTTTCACCTTGTTAAACCAATATCCAAAGTTCC 291
 |||
 QY 241 ttataattcttggagatcatcttgggctctgttttcttataattctggaagc 300
 |||
 DB 292 TTTATTTCTTTCAGAGATATCCATCTGGGCTCTGTTTGTCTTATTAATCTGAGGC 351
 |||
 QY 301 ttccataatgcagtgaaaaaaccacagaactctgggaatttgattacattgatg 360
 |||
 DB 352 TTCCTAATTCAGTGAAGAAACACAGAACTGTGATTAATATGAGCCGAATATAG 411
 |||
 QY 361 acctt 365
 |||
 DB 412 AATTT 416

RESULT 4
 LOCUS A1149899/C
 DEFINITION A1149899 508 bp mRNA EST 10-NOV-1998
 q14h06.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1752827
 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;
 mRNA sequence.
 A1149899
 A1149899.1 GI:3678368
 EST.

ACCESSION
 VERSION A1149899
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 508)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaído, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/db/ftp/image/image.html
 Insert Length: 742 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 455.

FEATURES
 source
 1..508
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1752827"
 /clone.lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: p7T73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCGAATGGAGGCGGCCCAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaído."

BASE COUNT 186 a 91 c 89 g 142 t
 ORIGIN

Query Match 47.9%; Score 215.4; DB 10; Length 508;
 Best Local Similarity 95.3%; Pred. No. 1.6e-38;
 Matches 222; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 133 agaaaaatgaataacttagagactatccagatccctgttggaaattagacctttctt 192
 |||
 DB 508 AGAAAAATGAATACTTAGGACTATCCAGATCCTTGGAAATATGACCTTTCTTT 449
 |||
 QY 193 ggaatattctcttttcaacctgttaaacacatccaaagtttcccttatattct 252
 |||
 DB 448 GGAATATCTCTTTCAGATTTCACTTTGTAACCAATACCAAGGTTCCCTTAATTTCT 389
 |||
 QY 253 tcaggaatccattctgggctctgttcttcaatattctggagccttccaaattgca 312
 |||
 DB 388 TCAGGATATCCATTCGGGCTCTGTTTGTTCATTAATTCGAGACCTTCCTAATGCA 329
 |||
 QY 313 gtgaaaaagaataaccacagaactctgggaattttgattacattgatgacctt 365
 |||
 DB 328 GTGAAAAAGAAAAACACAGAACTGTGATTAATATGAGCCGAATATATT 276

```

RESULT 5
A1002083/c 482 bp mRNA EST 27-AUG-1998
LOCUS ot38b02.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1619019
DEFINITION 3', mRNA sequence.
ACCESSION A1002083
VERSION A1002083.1 GI:3202120
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 482)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TUMOR Gene Index
UNPUBLISHED (1997)
CONTACT: Robert Strausberg, Ph.D.
EMAIL: cgapbs-remail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Insert Length: 598 Std Error: 0.00
Seq primer: -40m13 fwd. EF from Amersham
High quality sequence stop: 393.
Location/Qualifiers
1..482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1619019"
/clone_id="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAGTGAAGTGGAGCGCCGCAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 157 a 85 c 99 g 141 t
ORIGIN
Query Match 47.7%; Score 214.8; DB 10; Length 482;
Best Local Similarity 70.7%; Pred. No. 2.2e-38;
Matches 318; Conservative 0; Mismatches 97; Indels 35; Gaps 1;
QY 1 atgagtttaagcaccgacacagtcggtgttcttgatctccccaagaatcactcct 60
|||||
DB 450 atggatttaagcaccgacacagtcggtgttcttgatctccccaagaatcactcct 391
|||||
QY 61 tcaagatatgagtcacagaaacttcaagcaccgaccccttcaagcaccgacccctgca 120
|||||
DB 390 tcagatatgagtcacagaaacttcaagcaccgaccccttcaagcaccgacccctgca 331
|||||
QY 121 aaattattctgtagaanaaacttaaggaactacagatcgtttggaattatg 180
|||||
DB 330 aaattattctgtagaanaaacttaaggaactacagatcgtttggaattatg 271
|||||
QY 181 acccttcttcttgagatctctcttcaacctgttaaacatatacgaagttccc 240
|||||
DB 270 tcatattattctgagacccctcttaattgacagtaaaagaaacacacaaactctgac 211
|||||
QY 241 ttataattcttcagatatccatctcgggctcgtgttcttcatattcgcgagcc 300
|||||

```

```

DB 210 AAACATCATTTGTGTTATCTTCACCAAAATAGTCAAGTGAAGCGTGTACTGCTGTT 151
QY 301 ttccctaatgcagtgaaagaaacacagaaactcgggaatttattacatctgag 360
|||||
DB 150 CT-----tggcaattttgattatcatattg 126
|||||
QY 361 acttcagcatattgaattatcattctctgccttcttcaatttgggtgcactca 420
|||||
DB 125 ACTTCAGCATTTATGAAATTTATTCATTTCTCTGCTTCTCAATTTGGGGTGGCAGCA 66
|||||
QY 421 gaggaattgtgattggaacaatgttgtga 450
|||||
DB 65 GAGGATGTGATTTGTAACAATGTTGTTGA 36
|||||
RESULT 6
AA707529/c 281 bp mRNA EST 12-JAN-1999
LOCUS ah1a1a2.s1 Soares_testis_NHT Homo sapiens cDNA clone 1292062 3',
DEFINITION mRNA sequence.
ACCESSION AA707529
VERSION AA707529.1 GI:2717447
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 281)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TUMOR National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TUMOR Gene Index
UNPUBLISHED (1997)
CONTACT: Robert Strausberg, Ph.D.
EMAIL: cgapbs-remail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/image/image.html
Insert Length: 456 Std Error: 0.00
Seq primer: -40m13 fwd. EF from Amersham
High quality sequence stop: 218.
Location/Qualifiers
1..281
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1292062"
/clone_id="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTTCACATCTGAGTGAAGTGGAGCGCCGCAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 94 a 54 c 45 g 88 t
ORIGIN
Query Match 43.8%; Score 197.2; DB 10; Length 281;
Best Local Similarity 98.5%; Pred. No. 2.1e-34;
Matches 199; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 249 tcttcagatatccatctcgggctcgtgttcttcaattcgtgagccttccaat 308
|||||

```

Query Match	35.1%;	Score 157.8;	DB 10;	Length 415;
Best Local Similarity	98.8%;	Pred. No. 1.3e-25;		
Matches 159;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

```

FEATURES
source
Location/Qualifiers
1. .484
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BT1-ame-d-04-0-UI"
/clone_id="UI-R-BT1"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pUT73D-Pac (Pharmacia) with a modified
polylinker, Site_1: Not I; Site_2: Eco RI; The library
UI-R-BT1 is a subtracted library derived from a mixture of
the following tissues: hippocampus, thalamus, mid-brain,
medulla, corpus striatum, cerebral cortex and testis. For
a detailed description of the library from which this
clone was derived, please visit our web site at
ratest.eng.uiowa.edu. The subtraction has been previously
described in (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG_LIB=UI-R-BT1
TAG_TISSUE=testis
TAG_SEQ=ACGCAG"
165 a _98 c 96 g 125 t
ORIGIN

```

Query Match 32.4%; Score 145.8; DB 10; Length 484;
 Best Local Similarity 82.0%; Pred. No. 6.3e-23;
 Matches 168; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

DB 161 agatccgttggaattacacattcttcttgagggtatctctcttccattacattgttaa 220
 |||||
 DB 404 AGATCCGTGTTGGAAATCAATGAACTTCATGTTGGAGTTGTTCCCTTTCACACCTGGTAA 425
 |||||

DB 221 aacacatccaaagttcccttatattcttccagagatccattctgggctctgtt 280
 |||||
 DB 424 ACCATFACCAAGTTCCTTTATATTTCTCTCAGAGATCTCTTCGGGCTCTGCTT 365
 |||||

DB 281 tgttataatctcggagccttccctaattgcagtgagaaagaaacacagaactctgg 340
 |||||
 DB 364 TGTTCATTAACTCTGGACCTTTCGATTCGTGAAAGAAACACAGACACTCTGA 305
 |||||

DB 341 gaatttgattacattgattctt 365
 |||||
 DB 304 TAAAAATGAGCCAGCGGAGTAATTT 280
 |||||

RESULT 9
 BE638325/c 224 bp mRNA EST 28-AUG-2000
 LOCUS BE638325 mouse testis apoptosis subtraction cDNA library Mus
 DEFINITION EST00022 mouse testis apoptosis subtraction cDNA library Mus
 BE638325
 BE638325.1 GI:9932068
 EST.
 VERSION house mouse.
 KEYWORDS Mus musculus
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 224)
 JIANG, H., LI, L. and LU, G.
 TITLE Expressed sequence tags from mouse testis apoptosis subtraction
 cDNA library
 JOURNAL Unpublished (2000)
 COMMENT Contact: Jiang H
 Department of Pharmacology
 Hunan Medical University
 88 Xiangya road, Changsha, Hunan 410078, China
 Tel: 086-0731-4405036
 Fax: 086-0731-4497661
 Email: 1j12epublic.cs.hn.cn
 Seq primer: T7 Promoter Primer
 High quality sequence stop: 224.

FEATURES
 source location/Qualifiers
 1..224
 /organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="MTA22"
 /clone.lib="mouse testis apoptosis subtraction cDNA
 library"
 /sex="male"
 /tissue_type="testis"
 /cell_type="spermatogenesis cells"
 /dev_stage="adult"
 /lab_host="JM109"
 /note="Vector: pUCm-T"
 BASE COUNT 58 a 40 c 61 g 65 t
 ORIGIN

Query Match 28.0%; Score 126.2; DB 10; Length 224;
 Best Local Similarity 76.4%; Pred. No. 1.8e-18;
 Matches 155; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

DB 25 ccggtgttcctcgattccctccagaatcactgttcagagatgagtcacagaactt 84
 |||||

DB 204 CCTGTTCCTGCTGTTCCCTCCAGAGGTCACCACTCAAGAAATACCAACAGACTT 145
 |||||
 DB 85 tcagccagcacccttccaaactcaagcccttgcaaaaattatctgtagaaaatgtaa 144
 |||||
 DB 144 ACAGGCACAGCCTATAATCCCAAAACGAGTCAGAGATCATTTATTAAGAAATTTGAG 85
 |||||

DB 145 atcttaggaactatccagatcctgtttggaattatgaccttcttggagttacttc 204
 |||||
 DB 84 ACTTTAGGACACACCAAGTCTCTTGGAAATCAATGAACTTTCATTTGGAGTCAATTTTC 25
 |||||

DB 205 ctttccttgtaaacacata 227
 |||||
 DB 24 CTTTTCACCTTGTAACCCATA 2
 |||||

RESULT 10
 AA435988/c 382 bp mRNA EST 09-NOV-1997
 LOCUS AA435988 zuo3a08.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730742
 DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCTYME ANTIGEN CD20 ;,
 mRNA sequence.
 AA435988
 AA435988.1 GI:2140902
 EST.
 VERSION human.
 KEYWORDS Homo sapiens
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 382)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Krizman, D., Kucaba, F., Lacey, M., Le, N., Lemon, G., Merra, M., Martin
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 TITLE WashU-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson Rk
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Possible reversed clone: similarity on wrong strand
 Insert Length: 733 Sca Error: 0.00
 Seq primer: -41m3 fwd. EF from Amersham
 High quality sequence stop: 280.

FEATURES
 source location/Qualifiers
 1..382
 /organism="Homo sapiens"
 /db_xref="GDB:592743"
 /db_xref="taxon:9606"
 /clone="IMAGE:730742"
 /clone.lib="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc., and primed with a Not I - oligo(dT) primer [5',
 TGTTCACATCTGAGAGGAGGCGGCCCAATTTTATTTTATTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pRT3D vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bernaldo."

BASE COUNT 135 a 73 c 67 g 107 t
 ORIGIN

Query Match 25.1%; Score 113; DB 10; Length 382;
 Best Local Similarity 100.0%; Pred. No. 1.5e-15;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 338 tgggaatttgatcatgactgactcagcattatgaattatccttcctcgcct 397
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 150 TGGGAATTTTGATTCATGATGACTTTCAGCATTTATTAATTAATTCATTTCTCTGCTT 91
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 398 tctcaatttgggtgctcaccagagagattgattgtgaacaatgtgtga 450
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 90 TCTCAATTTTGGGGTCCACCTCAGAGATTGTGATGTGACACATGTGTGA 38
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 11
 AA416972 387 bp mRNA EST 09-NOV-1997
 LOCUS AA416972/c zt94h05.s1 Soares testis_NHT Homo sapiens cDNA clone IMAGE:730041
 DEFINITION 3' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
 mRNA sequence.
 ACCESSION AA416972.1 GI:2077080
 VERSION AA416972
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 387)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
 White,X., Wylie,T., Waterson,K., Stephens,M., Tan,F., Theising,B.,
 WashU-NCI human EST Project
 TITLE Unpublished (1997)
 JOURNAL Washington University School of Medicine
 COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 656 Std Error: 0.00
 Seq primer: 41m13 fwd. ET from Amersham.
 FEATURES
 source Location/Qualifiers
 1..387
 /organism="Homo sapiens"
 /db_xref="GDB:5926570"
 /db_xref="taxon:9606"
 /clone="IMAGE:730041"
 /clone_11b="Soares_testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories
 , Inc. and primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAGTGGAGCGGCCCAATTTTATTTTATTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization to Cc15, and was
 constructed by Bento Soares and M. Fatima Bonaudo."
 BASE COUNT 138 a 77 c 67 g 105 t
 ORIGIN
 Query Match 25.1%; Score 113; DB 10; Length 387;
 Best Local Similarity 100.0%; Pred. No. 1.5e-15;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 338 tgggaatttgatcatgactgactcagcattatgaattatccttcctcgcct 397
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 145 TGGGAATTTTGATTCATGATGACTTTCAGCATTTATTAATTAATTCATTTCTCTGCTT 86
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 398 tctcaatttgggtgctcaccagagagattgattgtgaacaatgtgtga 450
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 85 TCTCAATTTTGGGGTCCACCTCAGAGATTGTGATGTGACACATGTGTGA 33
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 12
 A0108532 538 bp DNA GSS 29-AUG-1998
 LOCUS A0108532 CIT-HSP-237967.TF CIT-HSP Homo sapiens genomic clone 237967, DNA
 DEFINITION sequence.
 ACCESSION A0108532.1 GI:3485222
 VERSION A0108532
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 538)
 AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
 Venter,J.C.
 TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
 Map Building
 JOURNAL Unpublished (1998)
 COMMENT Other_GSSs: CIT-HSP-237967.TR
 CONTACT Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
 Seq primer: M13-21
 Class: BAC ends.
 FEATURES
 source Location/Qualifiers
 1..538
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="237967"
 /clone_11b="CIT-HSP"
 /sex="Male"
 /cell_type="Sperm"
 /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
 HindIII"
 BASE COUNT 140 a 108 c 115 g 175 t
 ORIGIN
 Query Match 25.0%; Score 112.4; DB 13; Length 538;
 Best Local Similarity 99.1%; Pred. No. 1.9e-15;
 Matches 113; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 337 ctgggaatttgatcatgactgactcagcattatgaattatccttcctcgcct 396
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 168 CAGGGAATTTTGATTCATGATGACTTTCAGCATTTATTAATTAATTCATTTCTCTGCTT 227
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 397 tctcaatttgggtgctcaccagagagattgattgtgaacaatgtgtga 450
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 228 TTCTCAATTTTGGGGTCCACCTCAGAGATTGTGATGTGACACATGTGTGA 281
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 13
 AA758635 394 bp mRNA EST 29-DEC-1998
 LOCUS AA758635/c ah67b04.s1 Soares_testis_NHT Homo sapiens cDNA clone 1320655 3'
 DEFINITION similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;,
 mRNA sequence.
 ACCESSION AA758635.1 GI:2806498
 VERSION AA758635
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 394)
 NCBI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

FEATURES
 source
 1. 394
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="1320655"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer 15' TGTTACCAATCTGAAGTGGAGCGCCGCCCAATTTTCTTTTCTTTT 3'. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 133 a 73 c 67 g 121 t
 ORIGIN

Query Match 21.0%; Score 94.6; DB 10; Length 394;
 Best Local Similarity 88.0%; Pred. No. 2e-11;
 Matches 103; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 249 tcttcagagatccatctcgtggtcgttctgttcattatctggagccttcctaat 308
 394 TCTTTCAGGATATCCATTCGTGGGCTCTGTATGTATTAATTCGTGAGCCTTCCTAAT 335
 309 tcgaatgaagaagaacacacagaactctggaatttgattacatgatgacttt 365
 334 TGCAGTGAAGAAACACACAGAAACTCTGATATTAATTGAGCCGATTAATGAATTT 278

RESULT 14
 LOCUS BE638321 218 bp mRNA EST 28-AUG-2000
 DEFINITION EST000013 mouse testis apoptosis subtraction cDNA library Mus
 ACCESSION BE638321
 VERSION BE638321
 KEYWORDS EST.
 ORGANISM house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL Unpublished (2000)
 COMMENT Contact: Jiang H
 Department of Pharmacology

Human Medical University
 88 Xiangya road, Changsha, Hunan 410078, China
 Tel: 086-0731-4805036
 Fax: 086-0731-4497661
 Email: 1j12@public.cs.hn.cn
 Seq primer: 17 Promoter primer
 High quality sequence stop: 218.
 location/Qualifiers
 1. 218
 /organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="M1A13"
 /clone_lib="mouse testis apoptosis subtraction cDNA library"
 /sex="male"
 /tissue_type="testis"
 /cell_type="spermatogenesis cells"
 /dev_stage="adult"
 /lab_host="JM109"
 /note="Vector: pUCm-T"

BASE COUNT 61 a 39 c 46 g 72 t
 ORIGIN

Query Match 18.4%; Score 83; DB 10; Length 218;
 Best Local Similarity 79.7%; Pred. No. 8.7e-09;
 Matches 98; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 104 ctcaagcccttggaataattatctgtagaataaatacttaggactatccaga 163
 DB 198 CCCAAACCGAGTGCAGAGATCATTAATTAAGAAATTTGAGATCTTAGGACACCCAGA 139
 QY 164 tctcttggaattagaccttctcttggaagtatctcttcacacctgttaaac 223
 DB 138 TCGCTTTGGAATCATGAACCTTTTCATTTTGAGAGTCATTTCTTTCACCTTGGTAACC 79

QY 224 cat 226
 DB 78 CCT 76

RESULT 15
 LOCUS B86842 421 bp DNA GSS 09-APR-1999
 DEFINITION RPI11-26F10.TP RPI1-11 Homo sapiens genomic clone RPI1-11-26F10, DNA sequence.
 ACCESSION B86842
 VERSION B86842.1 GI:2927974
 KEYWORDS GSS.
 ORGANISM human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 421)
 AUTHORS Adams M.D., Rounsley S.D., Zhao S., Field C.E., Bass S., Linher K., Golden K., Berry K., Granger D., Sun E., Wible C., de Jong P. and Venter J.C.
 TITLE Use of BAC End Sequences for Sequence-Ready Map Building (1998)
 JOURNAL Unpublished (1998)
 COMMENT Other GSSs: RPI11-26F10.TV
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdamad@stirg.org
 Clones are derived from the human BAC library RPI1-11. For BAC library availability, please contact Pieter de Jong (pieter@jeon.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (info@resgen.com). BAC end search page:

http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html

Seq primer: SP6
Class: BAC ends.

FEATURES
source location/Qualifiers

1. 421
/organism="Homo sapiens"
/db_xref="GDB:7509729"
/db_xref="taxon:9606"
/clone="RPC1-11-26F10"
/clone_lib="RPC1-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC11 Human Male BAC Library"
BASE COUNT 124 a 95 c 71 g 131 t
ORIGIN

Query Match 15.7%; Score 70.6; DB 13; Length 421;
Best Local Similarity 94.8%; Pred. No. 4.6e-06;
Matches 73; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 148 ttaggactatccagatccctgttggaaatagaccttcttcttggagttatctcctt 207
|| |||||
DB 345 TTGCAGACTATCCAGATCCTGTTGGAATTATGACCTTCTTGTGAGCTATACCTCTT 404
QY 208 ttacacttggttaaac 224
|||||
DB 405 TTCACCTTGTGTAAC 421

Search completed: November 29, 2001, 00:18:42
Job time: 9879 sec

THIS PAGE BLANK (USPTO)